DEVELOPING A METHODOLOGY FOR COGNITIVE RESEARCH WITH SOCIALLY-HOUSED CHACMA BABOONS

Andrew McFall

In fulfilment of the requirements of a Masters Degree in Social Science.

Submitted: 31 March 2004

Supervisors: Prof S.P.Henzi and Dr L.Barrett

Examiners Copy

ACKNOWLEDGEMENTS

I gratefully acknowledge the friendship and assistance of Ms Rita Miljo, who hosted me at CARE and made this project possible, and thank her for her constant and severe criticism.
I gratefully acknowledge Prof. Peter Henzi and Dr Louise Barrett for their supervision, support and funding.
I thank Mrs Karen Pilling for logistical support and arrangements while at CARE.
I thank Mr Bennet Serane and his team of maintenance staff at CARE for their technical assistance during the research.
I thank the volunteers at CARE for assistance, and especially Ms Lynne McAlley for her fine photography.
I thank the National Research Foundation (NRF) for financial assistance, which made it possible to pursue this research.
I thank two anonymous reviewers for their much needed and extremely valuable comments and criticism on the initial version of this thesis.
I thank Ms Dina Frommert for constant support, advice and criticism.

CONTENTS

Index of figures and tables

ABSTRACT

Testing on laboratory-housed primates has long been the standard for research in cognitive psychology and other areas. As an alternative to this, a group of socially housed chacma baboons (*Papio hamadrayas ursinus*) at the Centre for Animal Rehabilitation and Education near Phalaborwa in Limpopo Province, South Africa, were the subjects for a set of basic cognitive tests. The purpose of the tests was to explore the importance of analogical reasoning by means of testing perceptual and conceptual skills in baboons. The main aim of this research is to investigate the degree to which captive but socially housed baboons are useful as experimental subjects, and to develop an apparatus and protocol to perform these tests *in situ* in the baboons' home cages. Five baboons were chosen as the subjects for experimentation. All subjects completed three groups of tasks to a criterion of at least 80% success over four successive experiments. The tasks tested baboons' discrimination ability between two coloured tiles, a reversal of that same discrimination task, and a simple match-to-sample task. As a result of time constraints, further tasks testing conceptual ability had to be abandoned. A record was kept of environmental and social factors that may have influenced the motivation of the subjects. The time taken to complete each experiment correlated in many cases with the number of distractions experienced by the subjects. There appeared to be no significant correlation between the score attained by a subject and the number of distractions experienced by that subject. The greater number of distractions experienced by the subjects was a result of the more engaged social world in which these baboons exist. Consequently, their motivation to perform repetitive cognitive tests was decreased, and needed to be countered in novel ways. An apparatus and a protocol for testing under these conditions were developed. Testing baboons' cognitive skills in these circumstances is both possible and desirable for ethical reasons, though the process takes longer than under laboratory conditions

1. INTRODUCTION

1.1 General introduction

The use of laboratory housed subjects for behavioural research raises a number of ethical and validity issues. While not as controversial an issue as medical or physiological research, laboratory research of behavioural theories is nonetheless not ideal when it comes to considerations of animal health and well-being. Nor can it provide insight into behavioural actions or cognition in the presence of conspecifics. Most subjects of behavioural tests are tested in isolation, whether they are housed socially or individually. This separation from conspecifics and their social context will ultimately have some effect on the outcomes of the experiments (Hare, 2002), particularly in the case of social animals such as baboons.

The observation of wild animals in their natural environment for the purposes of behavioural and cognitive research is vastly preferable to working with those same individuals in a laboratory setting, at least from ethical and ecological validity standpoints. The dilemma with the fieldwork approach lies in the necessity for certain tests to be performed under controlled, or relatively controlled conditions (Hare, 2002). One alternative approach to this hurdle is to utilise animals housed in a situation as close as possible to their natural one.

A very limited number of *in situ* behavioural experiments has been performed with primates housed in a social environment (Drea and Wallen, 1995; Custance *et al.*, 2002), so very little exists in the way of protocols to guide this research. A number of important questions therefore raise themselves. Firstly, is it at all possible to perform cognitive experiments with socially housed individuals in such an environment? Are the results that are generated comparable to those of more traditional experimental set-ups? And what pitfalls and stumbling blocks are inherent in this particular method of research? Hare (2002) argues that, to maximise experimental validity of a cognitive test with primates, experimental controls, however sensitive, are insufficient in isolation. Rather, a "combination of (1) proper control(s) and (2) an ecologically valid setting allows one to maximise the validity of cognitive experiments on primates" (Hare, 2002, pp270). External, or ecological, validity is achieved by directly relating the test design and desired behaviour of the subject to the behaviour of wild individuals as observed in the field (Custance *et al.*, 2002). The ecological validity decreases as the desired behaviour departs from the behaviour of wild conspecifics. The present study

attempts in some way to achieve that combination, primarily through the social setting of the experiments.

The venue for this research is the Centre for Animal Rehabilitation and Education (CARE), a baboon rescue facility in the Limpopo province of South Africa, located in an area known as Grietjie about forty kilometres from Phalaborwa. Baboons are raised in gender-balanced troops, with the ultimate aim of the centre being the release and reintroduction to the wild of the troop. Baboons are housed in large cages made of diamond mesh wire, with full view of their surrounds and of other baboons, both caged and wild. The floors to the cages are composed of natural earth (essential to the happiness of such fossorially oriented animals) and are supplemented by hanging swings and other items for them to play with. Cages are also built to incorporate existing vegetation, which is - at least initially - another plaything. The caged baboons are very well fed; it may be said that some are overfed, and are showing the effects. The inhabitants of the cages have ample opportunity for contact and interactions with members of their own cage community, adjacent troops in adjoining cages, and with wild baboons that are resident at the centre. Caged baboons at CARE are well fed, well cared for, well stimulated and generally happy primates.

Therein lie the difficulties for this research. Working with isolated individuals in an experimental enclosure (for example Bovet and Vauclair, 1998) allows the researcher to manipulate such desires and needs in the subject so as to produce a replicable set of results in a short period of time. The researcher can motivate the subject with food, attention, or other stimuli in order to perform the required task. These typical rewards are more or less denied to a researcher working in such an environment as CARE. Alternatives to this approach are therefore needed.

As the physical environment of the centre is rocky and inclined, cages have been built to suit the surrounds. This makes the use of traditional testing equipment impractical, and the construction of more suitable testing equipment was required.

A simple discrimination task was therefore used for comparison with laboratory trials. The task required of the baboons was well within their mental capabilities, and is well documented under laboratory conditions (see training tasks of Bovet and Vauclair, 1998; Depy *et al.*,

1999; Martin-Malivel and Fagot, 2001). The question then, is not whether the subjects were capable of performing the task, but rather how they performed under these conditions.

Once the subjects reached criterion for the discrimination task, they proceeded to more complex tasks. The match to sample task testing perceptual similarity matching required substantially more attention to successfully complete than did the simple learned discrimination task. One could predict that this necessary increase in attention would lead to an increase in the time taken by the subject to complete a set of trials. The intention was then to proceed from these basic tasks to more complex tasks aimed at testing analogical reasoning abilities, judging relations between pairs based on conceptual similarities. This aspect of experimentation was not achieved due to time constraints and experimental difficulties.

1.2 Literature review

The distinctive social skills of the primate order are commonly accepted to be associated with the increased brain size of the members of that order. This concept was originally known as the Machiavellian Intelligence Hypothesis (Byrne and Whiten, 1988), but is better known today as the 'social brain hypothesis', courtesy of Barton and Dunbar (1997), and Dunbar (1998). Amongst the primate order, however, there exist significant differences in competence both in general cognitive abilities, and in social behaviour and social cognitive skills.

The contrasts within the order seem to parallel differences between the species in key aspects of brain evolution (Rapoport, 1999). These contrasts are especially pronounced in the variation of the volume of the neocortex as a whole (Dunbar, 1992; Barton and Dunbar, 1997) and of the frontal lobe in particular (Dunbar 2002; Semendeferi *et al.*, 1997). They are evident to a lesser extent in the volume of certain key sub-cortical components associated with the processing of emotion – for example the amygdala, as shown by Emery and Perrett (2000).

Since its inception, the social brain hypothesis has had a major impact on the field of comparative primate cognition (Tomasello and Call, 1997). It is also of increasing importance to studies of human psychological adaptation as well as human developmental psychology (Whiten, 1991; Baron-Cohen, 1990). In respect to human developmental psychology, the primary differences between human and non-human primates appear to stem from the human capacity for metarepresentation and the use of a symbolic language system. However, despite extensive psychological experimentation, the exact nature of the cognitive differences

between monkeys and apes remain unclear, and the degree to which apes are capable of any form of metarepresentational thought - whether socially-based or not - remains contentious (O'Connell and Dunbar, 2002). This may be because the tests have not been exhaustive and/or the right kinds of tests have not been performed (Schrier and Thompson, 1984; Hare, 2002).

The only experimental study on monkeys' conceptual reasoning abilities is that of Dasser (1987). She attempted to test for understanding of relationships in macaques using a design in which monkeys were trained to match photographs of mothers with photographs of their infants and then tested on stimulus discrimination and match-to-sample trials. Dasser concluded that her study provided evidence that monkeys had some concept of the mother-offspring 'bond'. However, the tests did not actually probe whether individuals were able to judge relations between relations (meta-relationships). Animals were not asked to match mother-offspring pairs with other mother-offspring pairs, but merely to match a photo of a mother with one of two potential offspring, or to discriminate between a picture of a mother-offspring pair and a non-mother-offspring pair. This need not require conceptual knowledge, as the monkeys could have learned to discriminate the animals that appeared most similar physically, so that they succeeded in the tasks by using perceptual cues alone (see Thompson and Oden [2000] for a similar criticism of Burdyn and Thomas's [1984] study alleging that monkeys are capable of analogical reasoning). In this study, we exploit Dasser's basic design, but with critical modifications to ensure that animals are asked to judge the relations between stimuli and not simply to match them perceptually or learn simple discriminative cues.

Brain size differences between primate species are most pronounced in the region of the pre-frontal cortex. In humans many of our more advanced cognitive skills are associated with improved executive function and the expansion of this region, particularly the lateral prefrontal area. One skill that may be particularly relevant is the ability to think 'off-line'; that is, to think about objects and events that are not currently presenting a sensory stimulus. This ability allows us to reason analogically (eg, to think of time in terms of space: Boroditsky, 2000); to think through a complex chain of events to a final outcome; to compare a number of possible alternative outcomes to a given scenario; and, as Tulving (1983) puts it, to engage in 'mental time-travel'.

The prefrontal cortex is also associated with aspects of autonoetic consciousness: the sense of self, or the awareness that we are individuals who persist through time (Wheeler, 1997). The

capacity to view the world on a longer 'time-frame' may thus be the key to understanding exactly how monkeys, apes and humans differ from each other. That is, the differences between monkeys, apes and humans can be interpreted in terms of the depth of time frame over which a species' executive system allows it to manipulate its knowledge base in a virtual world.

It has also been suggested that, functionally, taxonomic differences in competences of this kind reflect differing patterns of demand generated by the level of social dispersion (Dunbar, 1998; 2002). It has been shown (Dunbar, 1998; 2002) that apes - including humans - lie on a separate grade to monkeys in terms of the relationship between group size and neocortex volume (with further possible grade distinctions within the monkeys). They have argued that this reflects a core difference between hominoids and simians in the extent to which they live in dispersed social system. That is, the more dispersed the social situation, the greater the requirements on the neocortex to facilitate social knowledge.

In this respect, recent work by Boroditsky (2000) arguing that (in children) the sense of time emerges by metaphorical analogy from a sense of space provides a promising basis for an explanation. Monkeys such as baboons have well established and profoundly competent mental mapping capacities (Byrne, 1995; Tomasello and Call, 1997), but seem to lack a clear sense of the future: baboons live very much in the here and now, with little evidence to suggest that they intentionally attempt to optimise behaviour over a time frame beyond today. However, following Boroditsky (2000), a sense of space might, with sufficient additional processing power, be used to develop an extended sense of time that could be factored into the prediction of future behaviour in a very effective way. If that step is scaffolded by analogical reasoning as Boroditsky (2000) suggests, then a plausible suggestion would be that, combined with causal reasoning and the ability to reflect on a virtual world (i.e. mentally rehearse future sequences of events), an extended sense of future time might provide the basis from which an ability to temporally sequence social events could emerge within the hominoids, leading eventually to full Theory of Mind (ToM) in humans. The core difference between humans, monkeys and apes may simply reflect the number of alternative future scenarios that can be managed and compared simultaneously. This will be a function of working memory capacity and the ability to interrelate temporarily-stored information across domains. Both the frontal lobe and the prefrontal cortex of the monkey are significantly smaller than those of humans and apes: 11% of total cortical volume, compared to 17% and 36% for chimpanzees and humans respectively) (Semendeferi *et al.*, 1997) and, since the

prefrontal cortex in particular is the brain area most closely tied to executive function and working memory, it suggests that monkeys may be more limited in this regard than apes and humans. By the same token, non-human apes will be more limited than humans.

It has also been suggested that the computational load placed on individuals living in dispersed social systems like the hominoids (where individuals see each other only at infrequent intervals, often weeks apart) is significantly greater than that when members of the group see each other every day (as is typically the case in monkey societies). The core to social survival in dispersed systems is firstly the ability to work with a virtual social world rather than one that is physically substantiated, and secondly the need to be able to project into an uncertain future key aspects of social interaction and relationships in a much more demanding way. If ToM (and related aspects of hominoid social cognition) is less about the ability to represent other minds than about the ability to represent the self in alternative spatio-temporal locations (Donald, 1991), then formal metarepresentation (as reflected in the ToM capacities that children develop around age 4 years) may simply be the scaffolded outcome of an earlier ability to represent the self in different spatio-temporal locations. In effect, episodic and autobiographical memory becomes the core to understanding advanced social cognition, although this does not discount the possibility that the ability to understand other minds may be critical to the final development in humans of ToM and higher orders of intentionality.

Because planning in effect involves future "time travel", an important component of this is likely to be the ability to inhibit prepotent responses. Humans (after a certain age) are especially competent at this (unless they have suffered damage to certain key brain areas, notably the prefrontal cortex: Bechara *et al.*, 1994). Like very young children, both chimpanzees and squirrel monkeys fail on these kinds of tasks unless either language-based symbolism (in the case of the chimpanzees: Boysen *et al.*, 1996) or an inhibitory intervention (squirrel monkeys: Anderson *et al.*, 2000) allows them to distance themselves from the immediacy of the reward. The ability to focus on the social domain may allow anthropoid primates to distance themselves sufficiently from the immediacy of the reward situation to reason analogically.

It has been argued, most notably by Dunbar (2000), that social cognitive abilities like ToM are more plausibly interpreted as the outcome of the ability to apply basic cognitive abilities like categorization, causal reasoning and analogical reasoning over an extended time frame in a virtual mental world (see also Harris, 2000). Hence, the differences between monkeys, apes

and humans can be interpreted in terms of the manner in which a species' executive system allows it to manipulate its knowledge base in a virtual world.

Although much has been made of ToM and its importance in human socio-cognitive development, our understanding of the cognitive processes involved in this important phenomenon has not advanced significantly in the decade or so since it was first proposed. We have learned a great deal about its natural history; but, as Roth and Leslie (1998) observe, we do not really understand exactly what this phenomenon is or why it emerges only at around 4 years of age in normal children - and not at all in autistic individuals (for a general review, see Barrett, Dunbar and Lycett, 2002).

The other central prediction of the social brain hypothesis is that, for primates, problems presented in a social context are more salient and easier to solve than equivalent tasks presented in abstract form or as ecological tasks. Although a great deal has been made of the pre-eminence of social intelligence (as opposed to "environmental intelligence") in primates, the evidence on which this claim rests (e.g. Cheney and Seyfarth, 1990) is, at best, inconclusive (Heyes, 1994; 1998). For example, it is not at all clear whether the original Cheney/Seyfarth experimental designs used to test social versus environmental knowledge compare like with like (i.e. are "fair" tests in the Popperian sense) or even whether their tests of environmental knowledge are ecologically valid for the species on which they were tested (Vervet monkeys, *Cercopithecus aethiops*). Drea and Wallen (1995) argue in a study of rhesus monkeys that rapid learning in previous experiments is more likely attributable to particular procedural or species differences between studies, rather than to social or ecological factors. In a major review of primate cognition, Tomasello and Call (1997) caution against over-interpretation of the available data, and argue instead that the data will support only an agnostic approach to the issue, as was initially proposed by McPhail (1985).

Some evidence that primates are sensitive to the emotional salience of social cues is given by the fact that, during conflict (but not other non-agonistic social interactions), baboons endeavour to keep the opponent in the left visual field, presumably in order to focus visual cues to the emotionally more sensitive right hemisphere so as to maximise speed of evaluation and response (Casperd, 1998). Since social stimuli are likely to have greater emotional valency than physical stimuli, this suggests that there may be some neurological basis for the predictions of the social brain hypothesis. Valid tests of human, ape and monkey problem-solving capacities are therefore crucial and will allow us to put the social brain hypothesis on a sound empirical footing.

There is significant experimental evidence in support of some of these basic elements. This includes: (a) birds have been shown to possess episodic memory, at least in rudimentary form (Clayton and Dickinson, 1998), (b) chimpanzees demonstrate evidence for causal reasoning in both physical (Limongelli *et al.*, 1995) and social domains (Dunbar, 2000) although monkeys fail (at least on physical reasoning tasks: Visalberghi and Limongelli, 1994; O'Connell 1995), (c) at least 2 species of great apes (chimpanzees and orang-utans) show evidence for the ability to mentally rehearse solutions to "problem box" tasks (ie, are able to open a puzzle box faster after being given the opportunity to reflect on possible solutions), although they do not do so as effectively as young children (Dunbar *et al.*, 2002), (d) that at least some great apes, for example gorillas, exhibit some evidence for the ability to plan physical tasks (Byrne, 1995) and (e) that adult humans differ by at least a factor of two in their ability to operate in an intentional domain (humans habitually manage fourth order tasks, whereas chimpanzees can at best cope with second order tasks: Kinderman, Dunbar and Bentall, 1998). These findings provide *prima facie* support for the direction that we propose to develop. More importantly, when known abilities to solve false-belief-type tasks for a monkey, ape and human are plotted against frontal lobe volume, they yield a clear linear relationship; interpolating frontal lobe volumes for other apes suggests that all the great apes lie close to each other, well separated from both lesser apes (gibbons) and monkeys (see Fig. 1.1).

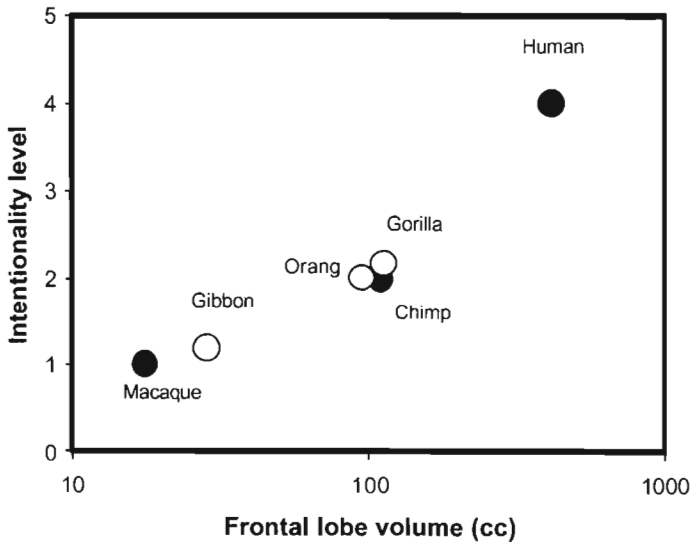

Fig. 1.1 Actual and predicted social cognition (indexed as levels of intentionality successfully achieved) plotted against volume of frontal lobe. Solid symbols: known levels of intentionality; open symbols: intentionality level predicted by regression equation for solid symbols.

Source: Dunbar (2002)

The key hypothesis, then, is that four cognitive mechanisms play a crucial role in fully functioning human adult social cognition and a progressively lesser role in the equivalent capacities of great apes and monkeys. These mechanisms are: (i) the ability to reason

analogically, (ii) the ability to reason causally, (iii) the ability to suppress prepotent responses in order to be able to evaluate the outcomes of alternative sequences of events and (iv) the ability to reflect on the first two processes in virtual format so as to evaluate the consequences of alternative courses of action.

1.3 Aims

The primary aim in the present study is to investigate the degree to which captive but socially housed baboons are useful as experimental subjects. The other main aim is to explore the importance of analogical reasoning, as an initial step towards a more comprehensive testing of the complete hypothesis.

The vast majority of other studies use animals that live either alone or in small atypical groups and animals are tested alone in a separate testing area. Few discrimination studies have been performed that attempt to integrate aspects of external validity such as testing individuals while in their normal social grouping and ideally with tasks that would form a normal part of that animal's behaviour (Drea and Wallen, 1995). Here, socially housed baboons were used and testing was performed *in situ* in the home cage. Such testing did indeed prove feasible and produced reliable results, and as such it should open the way for more comprehensive testing of other animals maintained under these conditions. It will also make comparisons between wild and captive animals more valid since, at present, there is always the possibility that an animal's failure or success in performing tasks is related to the captive situation *per se* and does not reflect the animal's natural abilities.

2. METHODOLOGY

2.1 Subjects

All subjects in this experiment were resident at the Centre for Animal Rehabilitation and Education (CARE), in Limpopo Province, South Africa. The centre is a baboon rescue facility housing in the region of 300 baboons at any one time, with the aim of releasing troops of baboons into the wild. There is also a small population of permanently resident baboons that are either unfit to be released due to illness or disability, or are animals that are otherwise unable to be integrated into a troop for release.

The subjects were composed of four females ranging in age from five to twenty-two years of age, and a male of approximately six years of age. The female subjects were Fiela, Julie, Tokoloshe, and Bobby (22 years old). The male was named Barnaby. The experiments were performed between December 2001 and September 2002. Photographs of each individual appear in Appendix 1. None of the subjects had ever been involved in behavioural experiments before the present study.

All subjects, with the exception of Bobby, were housed in the same cage complex. Bobby was housed in a cage that faced, and was in view of, the other subjects' cage complex. Of the study group, only Julie was housed in a completely social situation, in a large cage with seven other females and juvenile males. Tokoloshe was housed with a crippled baboon called Darwin. Bobby, Barnaby, and Fiela were in their own individual cages, each of which abutted other troop cages. This arrangement gave them physical and visual access through the cage wire to some of the other subjects, and other captive baboons.

For the purposes of this research, the term "socially-housed" refers broadly to baboons housed in a situation where they have contact with other baboons – visual, auditory and physical – at all times. Whether separated from other baboons by wire mesh or not, all the baboons in this study are aware of, and part of, an engaged social environment. At no time were they forcibly separated from this environment for the purposes of experimentation ie. those subjects housed in the same cage with other baboons were not separated from them for the purposes of research. All experiments were performed *in situ*. So while some of the study individuals were separately caged, they were still part of a broader social environment.

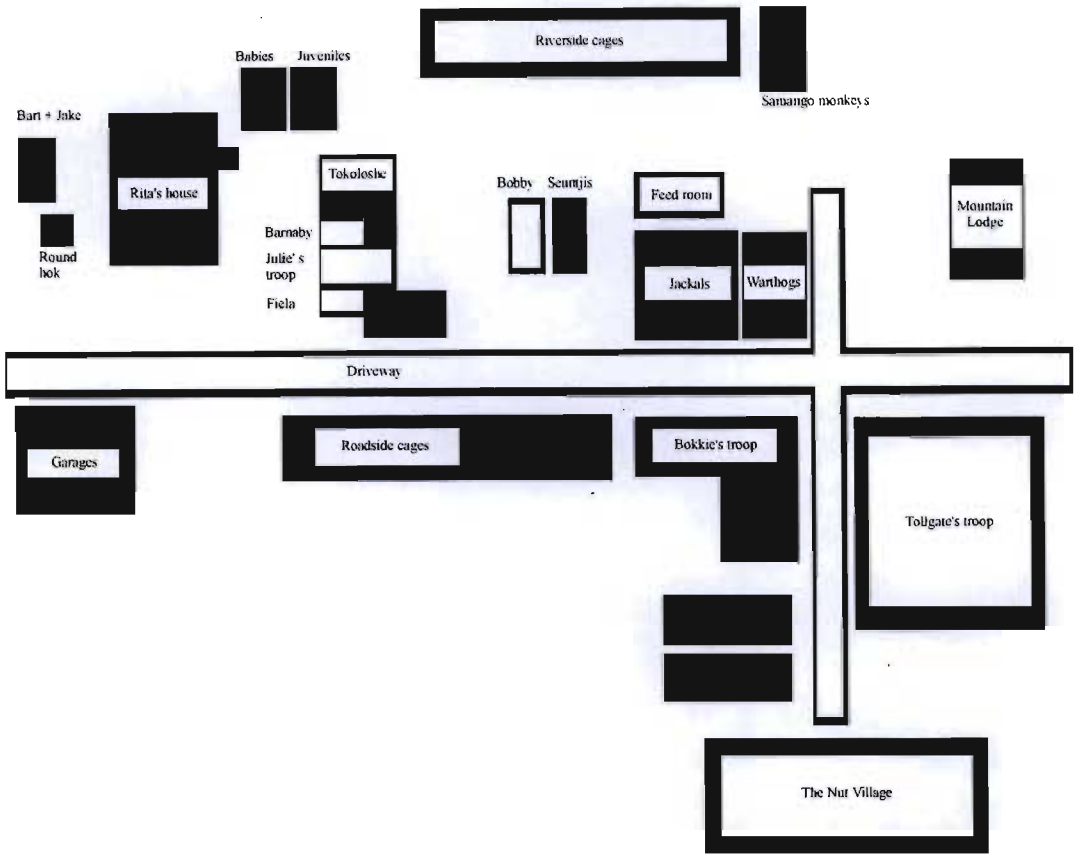

Figure 2.1: Plan diagram of the Centre for Animal Rehabilitation and Education.

2.2 Apparatus

The subjects were tested using a specially designed choice box. All tasks were performed using the same apparatus, with only the stimuli slides varying between tasks. The evolution and development of this apparatus is further discussed in Chapter 3.

The box was constructed out of wood, and comprised two separate compartments covered with hinged flaps. The flaps were designed so that the subject would be able to flip the lid of the box on the right only to the right, and similarly for the left hand side. The symbol stimuli were drawn on flat masonite tiles which were able to be slotted into the top of the flaps. The boxes were open to the experimenter to facilitate the introduction of the reward, without the subject being able to see or in any way detect which side the reward was on. Once baiting was complete, the experimenter could close this gap with hinged wooden gates, to prevent the subject reaching over and around the lid to extract the reward. The subjects were then required to open one of the lids, based on the presented stimuli on the respective flaps. If the

subject chose incorrectly, the experimenter was able to lock the lid of the baited compartment with a simple sliding lock mechanism. This denied the subject access to the food reward from the compartment.

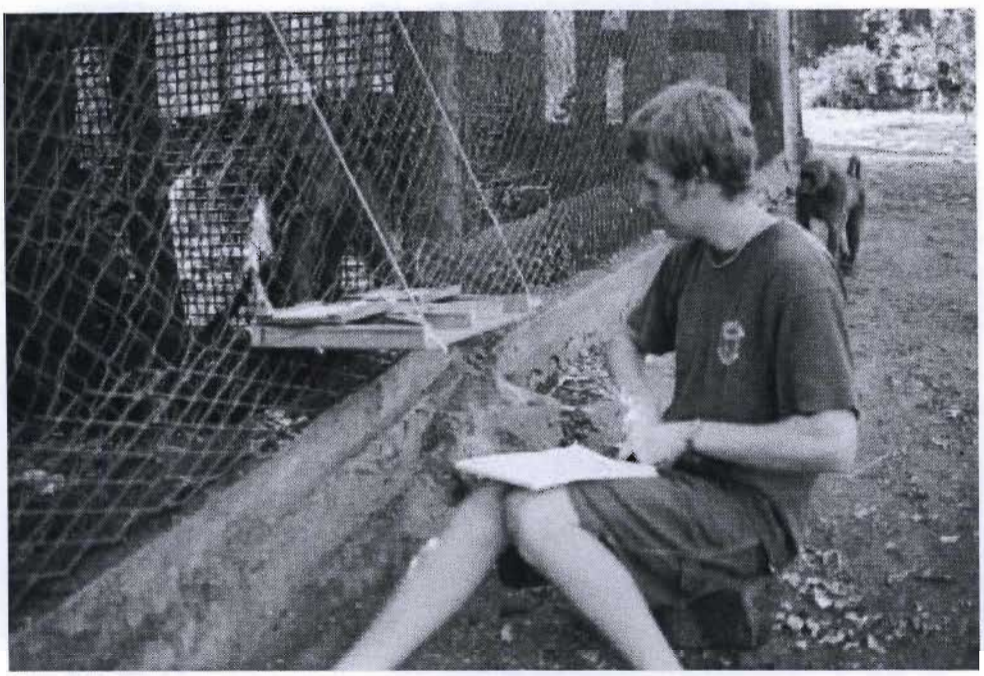

Figure 2.2: Apparatus in use with Barnaby. Wild baboon approaching at right.

The apparatus was attached to the subjects' cage by means of two cup-hooks mounted on the front of the box, and two suspension cords from the rear of the box. The hooks were connected to the wire, and the cords attached to the wire by means of small steel karabiners. In this way a level, though suspended, experimental platform was achieved. This was otherwise impossible given the slope of the land, and the rather *ad hoc* construction of the cages. The subjects gained access to the apparatus through two small, specially constructed holes in the wire of the cages. These allowed the subjects' arms through to open the lids of the box. A plan diagram and specifications for the apparatus appear in Appendix 2.

2.3 Experimental procedure

Subjects were initially trained in a simple discrimination task, followed by a reversal of the same discrimination cues. The discrimination required was on the basis of coloured tiles, one black and one white. Each experiment consisted of a set of ten trials with a subject before moving on to the next subject. The use of rewards – in the case of these tasks, little food items

- to reinforce a desired response is central to the concept of operant conditioning as related by B.F.Skinner (Http 1, 2004). The gaining of a food reward ("reinforcer") is contingent on the correct choice by the subject between two stimuli. In these tasks, there was no designed intermittency of reward delivery, one of the controls which may be used to increase the frequency of the desired response. The subjects would receive a food reward whenever they opened the correct lid of the apparatus based on the stimulus presented on that lid.

A variety of food rewards were used during the course of testing. The rewards consisted of foods that were not a normal part of the animals' diets. This was in an attempt to increase motivation through desire for the proffered food items. The most commonly utilised rewards were small marshmallows, which were absolutely silent upon introduction to the apparatus, and peanuts, which required more care on the part of the experimenter to minimise the effect of sound cues to the subject. Rewards also included other soft confectionary such as jelly tots, gum sweets, smarties and jube jubes. Varieties of nuts including almonds, cashews, brazils and hazel nuts were also used. Other rewards used were dried fruit, raisins, froot loops (breakfast cereals), acacia leaves, and, in the case of Bobby and Tokoloshe, biltong (dried, spiced beef). Interestingly, none of the other subjects shared their passion for biltong. This may be explained by the fact that both animals were kept as pets prior to their life at CARE, and they may have been previously introduced to it as a "treat" item.

To begin the process, the apparatus was attached to the wire of the cage in the correct position. The apparatus was situated at a suitable height that the subject could clearly see the discrimination stimuli, and would be able to open the lid of their choice without any hindrance. The position of the required tile on the box was varied at random after each trial, to minimise any success on the basis of a positional bias occurring. This was achieved by flipping a coin and recording the results – Heads=Left, Tails=Right – prior to the trials.

The white tile was then placed in the required position on the apparatus. Both lids were locked with the sliding bolt mechanisms to prevent them being opened prematurely. The food reward was then introduced. Prior to its introduction, the reward was hidden behind the experimenter's back, and manipulated into the hand corresponding to the required side, and hidden in the experimenter's fist. Both fists were then simultaneously brought into the subject's view, and placed in position in front of the hinged gates of the apparatus. In a swift

motion, both fists were unclenched and the gates pushed closed, in the process releasing the reward and pushing it into position.

The lids were then unlocked, the experimenter's hands removed, and the subject was allowed to open the lid of their choice. In the event of a correct choice, the subject was allowed to remove and consume the food reward. The process was then immediately repeated, once the white tile had been correctly positioned.

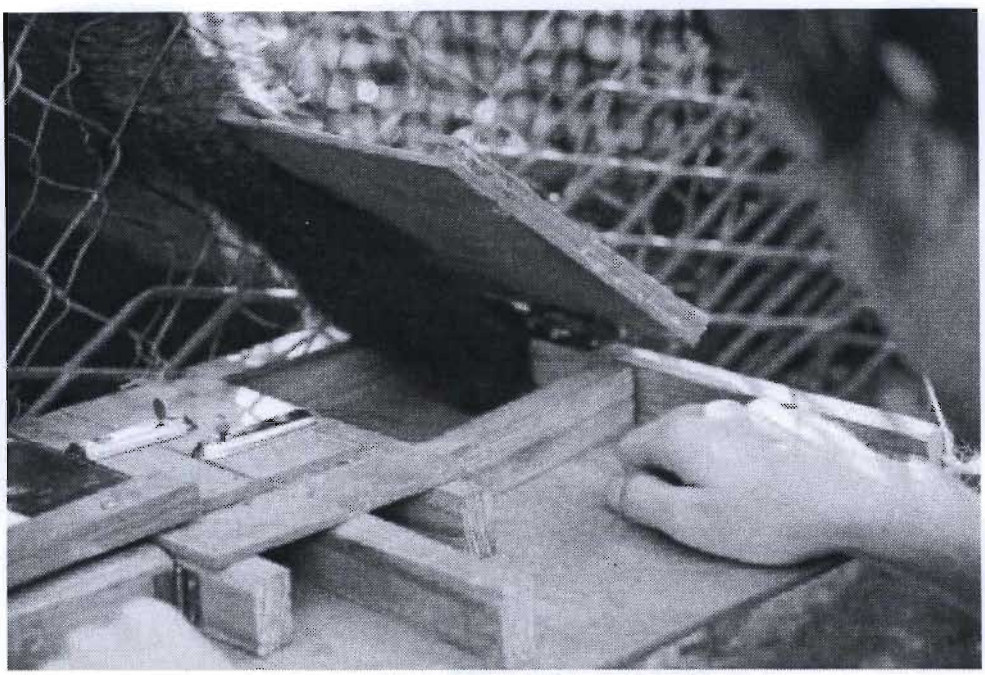

Figure 2.3: Barnaby removing food reward after correct choice.

In the event of an incorrect choice, the lid concealing the food reward was immediately locked, thus denying the subject access to the reward. The reward was then retrieved through the hinged gate at the rear of the box. The lid was then unlocked, and flipped open by the experimenter. A time-out period of one minute was then instituted. Both lids were left open to reinforce the absence of a reward during this period. Once the time-out period had expired, the white tile was correctly positioned for the next trial and the process repeated.

Figure 2.4: A time-out period after an incorrect choice, with both lids left open.

Once the required number of trials had been executed, the lids were again locked, and the apparatus removed from the wire. The subject would be given a few small titbits or a period of grooming time prior to the experimenter moving on to the next subject. This is in keeping with the Premack principle of using more probable behaviours to reinforce less probable behaviours (Http 2, 2004). In terms of this study, the less probable behaviours of discriminating between two available stimuli to gain small food rewards would be reinforced by the eventuality of receiving a larger food reward or period of grooming at the end of the experiment.

Once criterion of 80 percent success had been achieved over a period of four or more consecutive experiments, the reward position was reversed. The reward was instead placed under the black tile, and experiments were then continued until criterion under these conditions was achieved. As with the discrimination task, each experiment for this task consisted of ten trials with each subject.

The subsequent task was a match-to-sample (MTS) task, followed by a similar reversal (match-to-non-sample [MTNS]). For these tasks, twenty trials were performed with each subject to comprise an experiment, before moving on to the next subject. In the match-to-sample task, the subjects were presented with two differing tiles on the box lids. One lid had a

painted picture of a yellow banana, the other a painted picture of a green apple. These pictures were duplicated on separate tiles, which were only handled by the experimenter. Pictures of food items relevant to the subject were used in preference to abstract geometric shapes or numerals in a further effort to increase the ecological validity of the task.

The apparatus was prepared and positioned as for the discrimination and discrimination reversal tasks. Prior to baiting the apparatus with the food reward, and with both lids opened – thus hiding the tiles from the subject's view - the subject was presented with one or other of the experimenter's tiles for visual inspection. The sequence of presentations was determined on a random basis, again predetermined by flipping a coin. Heads=Left-hand tile, Tails=Right-hand tile. The tile positions were switched daily. The selected tile was held in direct view of the subject for a period of not less than 10 seconds, and no longer than 20 seconds. This tile was then removed from the view of the subject. The lids of the boxes were then closed and locked. The reward was then introduced, as previously described. The reward was placed under the lid with the picture corresponding to that which the experimenter had just presented. The lids were then unlocked and the subject allowed to choose and open one lid. The penalty for an incorrect choice was a "time-out" period of one minute and no reward, after which the trials would continue as normal.

The match-to-sample experiments were performed until each individual achieved a criterion of 80 percent success over a period of at least four experiments.

A match-to-non-sample task followed. The procedure was almost identical to the match-to-sample task. Once the subject had been presented with the tile, the reward was placed under the opposing box lid. If, for example, the subject was presented with the banana tile, the reward would be placed under the box lid covered by the apple tile. This set of experiments was continued to criterion of 80% success over a period of four experiments.

The intention was to then proceed to a more complex match to sample task, where pairs (ie. two identical banana pictures on one tile) of stimuli will be presented to the subject for inspection. The subject would then be presented with a choice between the same, paired stimuli and a non-identical pair of stimuli (ie. banana and apple picture). Once criterion was achieved in this task, a transfer task would be presented. This task would require the subject to use analogical reasoning to judge the relationship between pairs of items based on their

conceptual similarity. For example, the subject would be presented with a picture of identical paired stimuli (eg two apples), and then given a choice between a novel picture with a pair of stimuli (eg two oranges) and a novel non-identity picture (eg banana and sweet potato). The subject would be deemed to show conceptual understanding if they selected the identical pair (two oranges), since the relationship between the stimuli is the same as between the sample presentation stimuli. Unfortunately these tasks could not be performed with the subjects due to time constraints. As a result one of the aims of this study could not be achieved, seeing that this would have been a crucial step in investigating analogical reasoning and concept formation in baboons.

For each individual trial of an experiment, a number of measures were recorded. Firstly, the result of each trial was recorded – a simple correct/incorrect response. For each experiment of ten (or twenty for MTS and MTNS) trials a score index was created, simply counting the number of correct responses. In a discrimination trial, for example, with eight correct responses and two incorrect responses, the score was given as eight. To make comparison between the different tasks simpler, the score is sometimes presented as a percentage. Secondly, the time taken from the introduction of the reward into the experimental apparatus to the response of the subject was recorded. This measure was recorded using a simple digital watch (Xonix brand). Finally, comments were recorded regarding the attentiveness of the subject to the trial, and whether the subject experienced any distractions. A distraction was defined as any recognisable event or occurrence that drew the subject's gaze or attention away from the baited apparatus. Only the frequency of distractions was recorded – no attempt was made to ordinate the severity of distractions due to the disparate nature of these events.

For each entire experiment, a number of other measures were also recorded. The time at which each experiment was initiated and terminated was noted, to give an indication of the duration of each experiment. Any unusual weather conditions were also noted, for example extremes of heat or cold, precipitation or strong winds, which might affect the subject's willingness to play.

In the event of a subject not responding to the experimental stimulus in the presence of a reward for a lengthy period of time, a decision would be made to terminate the experiment. This time period was no less than ten and no more than fifteen minutes in duration. This was

necessary due to the limited windows of time available for effective experimentation. The possible reasons for non-response are discussed later.

The data were analysed using simple correlation and one-way ANOVA tests. A simple one-way ANOVA test was performed on the subjects' results from each task, and on the compiled results of all the individuals to compare their times, scores and distractibility with regard to the task performed. A two-tailed Pearson correlation was performed to compare the number of distractions with the time taken, and the number of distractions with the score for each experiment. These tests were to help ascertain the effectiveness of pursuing research under these conditions.

3. PILOT STUDY

The development of the apparatus and methodology was one of the primary aims of the current research, and the most fully realised. An account follows of the development of an effective apparatus to suit experimental conditions. A description of the conditions under which research took place along with some details of the social interactions of the subjects which affected experimentation, and the development of a reward system, is also included.

3.1 Apparatus design

The design of the testing apparatus evolved significantly through the process of experimentation. Further improvements to the eventual working design are both possible and necessary, and are shown in Appendix 2.

The initial design of a testing apparatus was that of two rectangular boxes attached by their bases to a large sheet of plywood, approximately 1200mm x 500mm. The boxes were open at the top, and closed with a non-attached lid. The lids had a small block of wood in the centre to allow them to be lifted. On the experimenter side of the boxes were small drawers, to allow the introduction of reward items unseen by the subject. No locking mechanisms for the lids were present.

The size and design of this apparatus meant that it needed to be deployed inside the cage, along with the experimenter. This was perfectly possible in the case of Fiela, who was the subject of some earlier pilot experiments. But it would not be possible with the majority of other individuals, due to either social housing or the unfamiliarity of most other baboons to humans, or both. The amount of contact with the experimenter could also influence the outcomes of the experiments, as the distractions of having a new playmate to groom and investigate can reduce motivation to work to roughly zero. Additionally, the process of manoeuvring the apparatus and the experimenter into the cage meant an increased risk of the subject escaping through the door.

The lids of the boxes proved problematic. Without a locking mechanism, there was no practical way of limiting access to the reward after an incorrect choice, short of forcefully holding the box lid down. This was sometimes successful, but always potentially hazardous,

in that there was the possibility of a confrontation with the subject over the disputed food reward.

The drawers to introduce rewards unseen were ineffective. The reward was visible for far too long, especially with the subject roaming all around, and not sitting obediently in front of the apparatus. The drawers were instead removed from the boxes, and shrouds of loose cloth were nailed around the openings at the base. This cloth was successful in concealing the introduction of the food reward, but was cumbersome and difficult to conveniently use.

A new design was therefore proposed and constructed. To allow the experimenter to remain outside the cage, and have the apparatus inside with the subject, it was proposed that an almost flat apparatus be constructed, with lids that flipped open away from the experimenter. It would also allow the subject to flip open the lid in a manner analogous to lifting a rock on the veld – picking up the side furthest from the body and peering under to provide a physical barrier to the possibility of a snake striking at them (were one to be present). This lid configuration would allow easy access to the locking mechanism while the apparatus was inside the cage, thus preventing the subject from receiving a reward each time, regardless of their choice.

The apparatus was designed to be inserted into the cage via a low, rectangular steel-reinforced trapdoor at floor level. Each trial would necessitate firstly the insertion of the apparatus into the cage, and secondly the removal of the apparatus once the subject had chosen which side to open. The removal of the apparatus would be necessary to change the position of the tiles on the lids, when required, and to re-bait the correct side of the apparatus prior to re-insertion.

For several reasons, this idea proved impractical. Firstly, and most compellingly, the Centre refused permission to cut the wire comprising the cage walls to attach the steel frame and trapdoor. Cutting the wire would compromise the integrity of the cage walls, and increase the likelihood of escape of the baboons. During the course of this study, it was observed that baboons are more than capable of breaking through triple layers of diamond mesh wire, particularly where the wire is attached to an immovable base. They do this by alternately pulling and pushing on the wire fronting where it is cemented to the floor, and eventually the repeated stress causes the wire to snap. Constant vigilance on the part of the staff is necessary to prevent escapes through this eventuality.

Secondly, placing the apparatus within the cage and in the grasp of the subject would have resulted in the swift dismemberment of the apparatus. Even with the eventual apparatus design being situated outside of the wire, and being of only limited access to the subject, several of the subjects managed to break parts of the apparatus off. This usually occurred when the subject was frustrated at being denied the reward after an incorrect choice, or when the experimenter had to leave the area momentarily to fend off wild baboons or break up a fight between them and caged baboons.

Finally, the repeated insertion and removal of the apparatus into the cage over the course of ten or twenty trials would have been both distracting to the subject, and exhausting to the experimenter.

Since the flattened design had already been constructed, it was decided to modify the design to allow the cage to be mounted on the wire forming the cage front. It was first proposed that a platform be built outside and abutting each cage, to support the apparatus for experimentation. This idea was modified to utilise the apparatus itself as a platform. This added to the mobility of the apparatus, and allowed it to be deployed on any wire-fronted structure with a minimum of difficulty.

The modifications were simple. Two cup-hooks were screwed into the front of the apparatus, with their points facing towards the ground. Two pieces of braided nylon cord approximately 80 cm long were attached with screws to the opposite side of the apparatus. Their loose ends were tied into separate loops, through which small steel karabiners were attached. The apparatus was attached by hooking the cup-hooks through the wire, and clipping the karabiners to the wire above the apparatus, thus suspending the apparatus in a level configuration. This arrangement was additionally very stable. The position of the lids was altered so that the lids opened laterally – still allowing the subject to cautiously flip open the lid as if they were lifting a rock. A clearer distinction between opening the left and right side was also achieved by this arrangement.

The subjects were able to access the apparatus through two specially cut and reinforced holes in the wire fronting the cage. Similar holes were in place already on many of the cages – particularly those of large, individually housed males – to facilitate feeding. This arrangement

was therefore acceptable to the Centre management. The two adjoining holes were large enough to permit the thoroughfare of the subject's arm up to the shoulder – more than sufficient for access to the apparatus. Many of the subjects bypassed the holes altogether, and were able to reach through the wire to access the apparatus and rewards.

The difficulty of insertion of the food reward for each trial was dealt with by having a hinged rear gate to each box. Two pieces of wood were attached to the centre strut of the apparatus by means of hinges. The hinges allowed movement of the gate much in the manner of a pinball flipper. To insert the food item, the gates were pulled back simultaneously with both hands, the food item dropped in, and the gates simultaneously closed. This allowed the insertion of the food item unseen, and mostly unheard, by the subject.

The surfaces of both the lids were spraypainted black. The insertion of a white painted tile on one side left the other side displaying only a black surface. The white tile was moved as necessary for the original discrimination experiments. Later match-to-sample experiments utilised two tiles with differing images painted on them. The tiles were slid into place under overhanging strips of masonite on either side of the lid.

Throughout the construction, all joints were both glued into place, and then reinforced by a number of screws to form a solid, rigid apparatus. This construction survived most attempts at dismantlement, but did at times fail. The baboons appeared to take great pleasure in dismantling anything that came within reach. Any small irregularity in the construction of the apparatus was first investigated with a fingernail, graduating to the fingers if they could get a positive grip, and followed by the physical removal of the piece if possible. The strips of masonite holding the tiles in place were obvious starting points. These were most often used to open the chosen lid in each trial. However, if the lid was locked, they would still attempt to open the lid to access the food reward, sometimes resulting in the strip being torn off. Due to the limited tolerances at the edge of the lid, these strips were simply glued together with wood glue, and reinforced with small wood screws. Very often it was not the joins that gave way, but the wood itself. The masonite would delaminate at the point where the screws ended, often right next to the layer of glue, which would maintain its integrity. Each time this occurred, the broken strips were salvaged if possible, and re-glued into place, and more screws were added to hold it together.

Another weak point was that of the hinges that held the lids to the main frame of the apparatus. Each side of the hinge was held down by three large wood screws, both screwed and glued into place. As the lid provided a large surface on which to pull, the hinges were vulnerable to breakage. Surprisingly, this only occurred a few times over the course of the year. The hinges of the gates at the back of the box were relatively inaccessible to the subjects, though Barnaby managed to bend them a little by pulling up on the gates in attempts to access a food reward in a locked compartment once or twice.

The only other areas of wear on the apparatus were the cup-hooks. The subjects sometimes grabbed the rear edge of the apparatus and pulled it towards themselves, to presumably gain a better look at it, or see if any food was hidden in it. This could be prevented when the experimenter was present by holding it down. In the momentary absence of the experimenter, it would mean that the cup-hooks would be bent straight or broken, necessitating replacement. This was a very minor problem, however, as repair was as simple as screwing in a new set and continuing. The plastic coating of some sets of cup-hooks came in for some attention, and was swiftly gnawed off, mostly by Tokoloshe.

The basic design of this apparatus is sound. The suspended platform configuration worked very well, as did the lid design and hinged gates at the rear. It is a practical, highly mobile, and functional apparatus, well suited to the conditions in which it was deployed. Naturally, there are several immediate improvements to the design that could be made. The first is that the depth of the boxes could be increased; just enough to allow the entrance of a closed human hand while the lid is closed. This would allow the introduction of the food reward directly into the box, with an absolute minimum of sound or visual cues. The closure of the gates sometimes sent the food reward tumbling into the box, which was very audible, followed by the immediate opening of that lid by the subject as a result. This effect was particularly noticeable with food items such as peanuts and hard sweets.

The issue of sound cues as a result of the materials used in construction was problematic for much of the research. Attempts were made to use only 'quiet' food items such as marshmallows and other soft sweets. The subjects quickly tired of these, to the point where they would refuse them if offered, and their motivation to play suffered as a result. A solution of a sort was found; a thick lining of felt was glued to the inner sides of the boxes, to muffle the impact sounds of harder food items. This worked very well. The disadvantage was that the

subjects immediately tried to groom this new substance, and found it rather easy to rip up with their fingernails, so constant vigilance on the part of the experimenter was necessary to prevent this occurring. A large supply of replacement felt was also necessary. The liberal application of glue to attach the felt limited the possibilities of the subject ripping the felt up, though it did tend to slightly negate the intended advantage of the felt by hardening it and thus increasing its sound transmitting properties.

The tiles were positioned and held in place by the strips of overhanging masonite. Over time the masonite of the tiles wore smooth, causing the tiles to fall out of the apparatus. This was dealt with by simply coating the base of the tiles with wood glue to add bulk, thus effecting a tighter fit with the lids. Another solution would be to add small catches to the open sides of the lid to prevent the tiles sliding out. This was deemed unnecessary at the time of the experiments as the glue solution worked perfectly. Whichever method of securing the tiles in place is used, it is essential that the tiles are able to be easily removed by the experimenter to allow their positioning.

3.2 Distractions

CARE is a working rehabilitation centre. Consequently, there is a large amount of human activity in and around the baboons cages. This activity provides a source of constant interest to the inhabitants of the cages – obviously including the subjects of the experiments. The timing of the experiments is thus of great importance to minimise the number of avoidable disturbances. The general schedule of work for the maintenance staff would be as follows. The staff would arrive for work at about 7am, whereupon they would begin changing and replenishing the water in the cages. The babies and juveniles would be served with their morning milk bottles, and the babies would be transferred from their inside sleeping cages to the large outdoor play cages (adjacent to the juveniles). Certain select older individuals would also receive a bowl of milk between 7am and 8am. These individuals included Fiela, Tokoloshe, and Bobby, and some other baboons which were recovering from illness and needed feeding up. Those individuals which received milk would become very agitated and excited upon the first sighting of anyone bearing a bowl. While the milk feeding was in progress it was impossible to work with those individuals, though once they had received their milk they were quite happy and motivated to play. The watering process would continue to around 9am. The staff would break for breakfast/tea between 9am and 10am.

When they returned from tea at 10am, the morning feeding would begin. This would be heralded by great excitement from the majority of the caged baboons, particularly when the staff would pass by their cages *en route* to their destination with shopping baskets or sacks full of food. Working with any individual during or around feeding times would be nearly impossible, as their attention would be on the purveyors of food, and not on the task at hand, regardless of the quality of the rewards offered. Also, once the animals had been fed, their motivation to work for small tidbits was noticeably reduced, if not non-existent.

Once the morning feeding was completed, the staff would then proceed on to general maintenance and repair work to the cages. New structures were regularly being erected – depending on the state of finances at the centre – to house the ever-growing population of baboons. The baboons appeared to regard this as background activity, and generally paid little attention to these goings-on. Often, though, the structures housing troops or individuals required repair, which elicited more attention from the subjects. On a number of occasions Barnaby broke through the wire separating him from Julie's troop. This would occur when Julie was swollen, and would be with the obvious intention of consorting. Under these conditions any work with either individual was absolutely impossible, and was suspended until such time as they were separated and the cage repaired.

The staff would break for lunch between 1pm and 2pm. Once they returned from lunch, the afternoon feeding would commence (depending on the availability of food which was obtained from local farmers), with the same effect on the baboons' attention spans and motivation to play. Feeding would continue until around 4pm, or sometimes 4:30pm. Maintenance would then continue until between 5pm and 6pm. The second milk feeding for both juveniles and selected adults would take place just before darkness fell. The babies would be brought inside to their sleeping cages. The staff would then depart for home.

As is apparent from this schedule, it was generally counterproductive to work during or near to feeding times, or major periods of activity. The windows of time open to experimental activity were therefore limited to only a few periods in the day. These were further limited by other factors such as the weather, unexpected visitors at the centre, and unannounced supplemental feedings of maize when it was available.

The first experimental period was therefore between 8am and 10am – after milk time, and before the morning feeding time. This period had the added advantage of minimal staff interference due to their tea-break between 9am and 10am. This period was also ideal as the animals had not been fed yet (though food was generally still present in their cages from the previous afternoon's feeding), and the subjects' motivation levels – prompted by hunger for the rewards - were higher.

In winter however, this was also the coldest part of the daylight hours. Subjects were often more intent on finding a little patch of sunlight to warm up after a cold night than playing with the experimenter. Subjects could be coerced into playing, though usually with some reluctance. Sometimes though, no response was forthcoming from the subjects, particularly on days of extreme cold, and especially when the cold was accompanied by a strong wind. On these occasions the baboons would simply sit on their hands and feet and look miserable. Those housed together as troops would huddle together for warmth. In summer, this was the ideal working period – especially for the experimenter - as the day was not yet too warm as to be uncomfortable.

The second experimental period was varied according to the season. It had to be timed so as to be at a reasonable interval after the last feeding, while not continuing into the next feeding period, nor continuing into the night. In winter, the most suitable time was between midday and 2pm. This was approximately two hours past the morning feeding time, and just prior to the commencement of afternoon feeding. The subjects would be comfortable and warm, and generally amenable to playing.

The same time period would not be available in summer, due to the extremely high temperatures at midday. Temperatures in summer averaged between 35 and 40 degrees Celsius. The subjects would spend this time doing as little as possible or sleeping so as to minimise overheating. Working during this period was also extremely unpleasant for the experimenter. The alternative time period was between 5pm and 7pm in the evening, when temperatures decreased to the low 30's, and the cages were out of direct sunlight.

This time period was also approximately one to two hours after the afternoon feeding. Difficulties arose if the food consignment arrived late in the afternoon, as feeding would commence immediately the food arrived. The experimental period would then be very close

to, or overlap with the feeding time, with a consequent drop-off in attention from the subjects. As is evident, it was extremely important to be flexible with experimenting times.

Of the subjects, different individuals were affected to varying extents by the weather due to differing morphologies. Tokoloshe, Julie and Barnaby originated from the northern areas of South Africa, and thus were well suited to the temperatures of Limpopo province. They were characterised by longer limbs and bodies, and sparse short, coarse fur which was light in colour. They were very similar in appearance to the wild baboons that occupied the area. Fiela and Bobby, on the other hand, originated from the Cape Province, and possessed the characteristics of baboons from those areas; shorter, stockier bodies and legs, and a thick coat of longer, finer, and darker fur. They also had darker sections of fur on the forearms and forelegs, which gave them the appearance of wearing dark socks. Their thick coats must have made the high temperatures doubly uncomfortable, though providing more warmth in the winter months. They also had rounder faces and proportionately shorter muzzles than their northern counterparts.

Throughout the experimental periods, the subjects were at all times free to interact with other individuals, should they choose to do so. These interactions were either with other inhabitants of the same cage, inhabitants of neighbouring cages, or with the wild baboons (aka 'the longtits') which occupied the area, and spent a great deal of time and energy trying to extract food items from the cages.

Of the subjects, only Tokoloshe and Julie shared cage space with other baboons. Tokoloshe was housed with Darwin, a crippled sub-adult male. Julie was the matriarch of a cage of seven other baboons. Both were the dominant individuals in their cages. Neither individual would allow any other occupant of the cage anywhere near the experimenter or apparatus. When they attempted to do so, the subject would physically attack, or chase away the interloper. Sometimes the subject would not even need that excuse. In particular, when Julie opened the incorrect side and received no reward, she would often turn her frustrations on the nearest target at hand and deliver a beating to that individual, or chase them around the cage. Apart from these interactions, there was little interference by other cage inhabitants with the experimental process. The only exception to this was when Jessica, a very low-ranking female in Julie's cage gave birth around the time when Julie was progressing to the match-to-non-sample task. Julie immediately lost all interest in the task, regardless of the quality of the

rewards, and took to following Jessica around and attempting to groom the baby. Experimentation was therefore terminated at this stage, as it would be fruitless to attempt to continue in the face of such competition.

Interactions with neighbouring individuals were generally antagonistic in nature. The presence of the experimenter seemed to escalate the intensity of these conflicts, as the subject would initiate a fight and then look to the experimenter for coalitional support. This was evident by the nature of the attack. The subject would approach their opponent while screaming and looking back to the experimenter, and then at the opponent, and repeating the process. Fiela in particular was fond of this approach when dealing with Julie. Tokoloshe also used it on occasion when fighting with Three-tits Bobby (a long-term neighbour). The one subject not using this approach was Barnaby, the only male in the study. He would simply fight with the nearest male – Kobus – with no regard for the experimenter. While Julie used it on occasion, it was more common for her to approach one of the two males (Barnaby or Kobus) for support against the experimenter, again when she had opened the incorrect side and received no reward.

However, the position of the apparatus in relation to the neighbouring cages was a point of concern. In the initial set of experiments with Fiela, the apparatus was attached to her cage approximately one metre away from the wire separating her from Julie's troop. Julie would position herself against the wire and watch the proceedings. Under theses conditions, Fiela would consistently open the side of the apparatus furthest from Julie, regardless of the position of the white tile. When the position of the apparatus was changed to being equidistant from both Julie's cage and the other adjoining troop, this condition ceased. In this position, Fiela also sat with her back towards Julie's cage, which may have further decreased the intimidatory effect of Julie. The position of the apparatus with the other subjects was then assessed, and found to be satisfactory in all instances.

Other than these interactions, there was no real consistent interference by the neighbouring baboons. They were always interested in the goings-on in the area – particularly since they involved food items – and would be an attentive audience. The large troop adjoining Fiela and Julie's cages would sometimes attempt to start a fracas with Fiela, when they noticed the volume of rewards Fiela was receiving. They were separated from Fiela by both wire and

sheet metal, obviously as a result of the frequency of their fights with her. Fortunately the sheet metal provided a good deal of protection from their attentions.

Of more concern to the experimenter were the wild baboons. The presence of food rewards on the experimenter was a strong attractant to the juveniles in particular. The adults tended to be a little more cautious in the presence of the experimenter. When a wild baboon approached the experimenter, the subjects would become very agitated and aggressive toward the interloper, in an attempt to drive them off and preserve the food resource for themselves. Barnaby was especially vociferous in this regard, also becoming very aggressive towards any other humans approaching the experimenter. In these instances trials would only continue once the offending party had departed to an acceptable distance away.

The presence and actions of the wild baboons sometimes provided other distractions to the subjects. The presence of a mother with a young baby was an irresistible attractant to any of the female subjects. Depending on the proximity of the pair, trials could usually proceed, with attention by the subject limited to the time-out periods. The wild females were also very wary of humans, so would not approach too closely, particularly when carrying a very young baby.

The noise generated by a fight anywhere in the vicinity would be sure to draw the attention of a subject. If the fight was in a neighbouring cage, the subject would run to the wire closest to the fight and join in with the fracas until it died down, or was broken up by a member of staff. This was only a temporary distraction, and trials could continue once it had died down. An alarm call, by either the wild baboons or by caged individuals, would be a little more of a distraction, as the subject would immediately climb to the top of their cage to see what was happening. Once they had returned to the ground, trials could continue. However, when there was a strong wind blowing, the incidence of alarm calls (usually false) increased considerably, disrupting the experimental process to a much greater extent.

Other weather conditions which detracted from the subjects' motivation to work included periods of rainfall, and extreme heat or cold. During rainy periods – which were very infrequent – the subjects simply sat under cover and looked miserable. Much the same behaviour occurred during periods of extreme cold and wind, with the difference that the subjects attempted to make best use of any patches of sunlight, if any was present. The

extreme heat of summer provoked listlessness and as little movement as possible, which meant little attention was paid to the task, and given rather to finding the shadiest spot.

The reproductive state of the female subjects – barring Bobby, who was post-menopausal - was of some significance to their motivational state. While the state of the genital swelling had no observable inhibitory effect on the females' motivation to work, the presence of the male experimenter apparently did. As the females approached their fully swollen state, they presented their posteriors to the experimenter with increasing frequency. The experimenter had to pause between each trial to scratch the side of the subject's presented bottom to encourage them to turn around and continue with the next trial. This constituted no more than a nuisance factor, as did their increased predilection for grooming the experimenter.

This behaviour became dangerous, though, in the presence of adult wild males. Some of the wild baboons adopted a cage complex as their own 'troop', and objected very strongly to any other male – including human males - interfering with 'their' females. The male who had claimed ownership of the experimental cage complex was named Tarzan – a young, well-built baboon who was challenging the alpha male (Tollgate) for primacy of the wild troop. On several occasions Tarzan became extremely threatening towards the experimenter, on at least one occasion attempting an attack. Fortunately this attack was thwarted by the experimenter, though it was a close call. Thereafter it was deemed necessary for the experimenter to carry some form of weapon – a stout stick sufficed - in case of future aggression. This was prudent, as Julie sometimes attempted to solicit support from Tarzan when she was denied a reward after an incorrect choice. Brandishing the stick or throwing a stone generally warded him off. However, this possessive behaviour by adult males – both by wild baboons and experimental subjects - should be borne in mind by future researchers with a care for their own personal safety.

3.3 Determining rewards

The relationship between motivation to work and the desirability of the reward on offer is one of great importance under these conditions. Since the subject has the option of not working if they don't want to, the desirability of the reward to the subject must be greater than that of a reward used under a laboratory situation, where the subject is typically deprived of food for a period before the initiation of the trials (eg. Menzel and Juno, 1984). With the design of the

experimental apparatus, the choice of reward was further dictated by the physical make-up of the reward and its propensity to create a noise on introduction to the apparatus. So, given the two constraints of desirability and sound qualities, it is evident that peanuts, for example, score low on both counts. The baboons are sometimes fed peanuts, so they are not really a special treat, and the hardness of the peanut created a great deal of noise when introduced into the apparatus. Marshmallows were a very 'quiet' reward, but repeated exposure to them as rewards lowered their desirability to the subjects, to the point where Barnaby, for example, would not even approach the apparatus if he saw that the 'reward of the day' was marshmallows. He would accept them only if they were interspersed with jelly tots or smarties, to provide some variety, and then with some reluctance.

The discovery that Tokoloshe found biltong irresistible came quite by accident. The experimenter was eating some biltong while in the vicinity and Tokoloshe begged a taste. Once she tasted it she immediately ran to drink some water (biltong is very salty) and then returned for more. Biltong was then utilised as a reward item with Tokoloshe and Bobby, as none of the other subjects would accept it. It is evident from the results that Tokoloshe was highly motivated to work, and proceeded through to the match-to-non-sample task. None of the other subjects reached this stage.

It was important, for all of the subjects, that the rewards offered by the experimenter were varied constantly, so as to avoid boredom with the rewards. Other rewards could be used to complement the food rewards, such as allowing the subject to groom the experimenter for a period of time in the advent of a correct response, or at the end of an experimental session.

4. RESULTS

As an initial overview of the results, the means are presented for the number of trials required by a subject to reach criterion of 80% success over four consecutive experiments (40 trials for discrimination and discrimination reversal tasks, and 80 trials for the match to sample task). For the discrimination task the mean equals 466 (range=210-610), for the discrimination reversal task the mean equals 468 (range=90-720), and for the match to sample task the mean equals 1332 (range=900-1500). This indicates that the subjects found the match to sample task more complicated than the first two tasks, and had difficulties in performing it, or lacked the motivation to tackle the task as it lacked the novelty value of the initial ones.

As mentioned previously, for each experiment of ten (or twenty for MTS and MTNS) trials a score index was created, simply counting the number of correct responses. Due to the differing numbers of trials comprising an experiment, and to facilitate comparison between the different tasks the score is presented as a percentage in the following graphs.

4.1 Discrimination learning

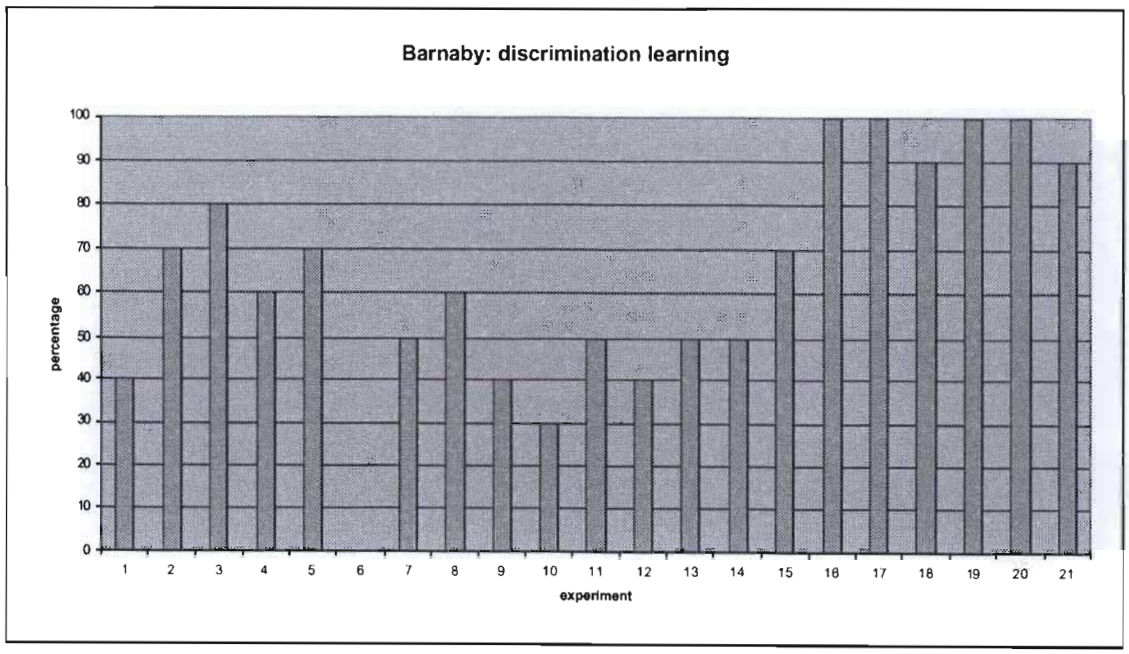

Figure 4.1: Percentage scores achieved by Barnaby during discrimination learning experiments.

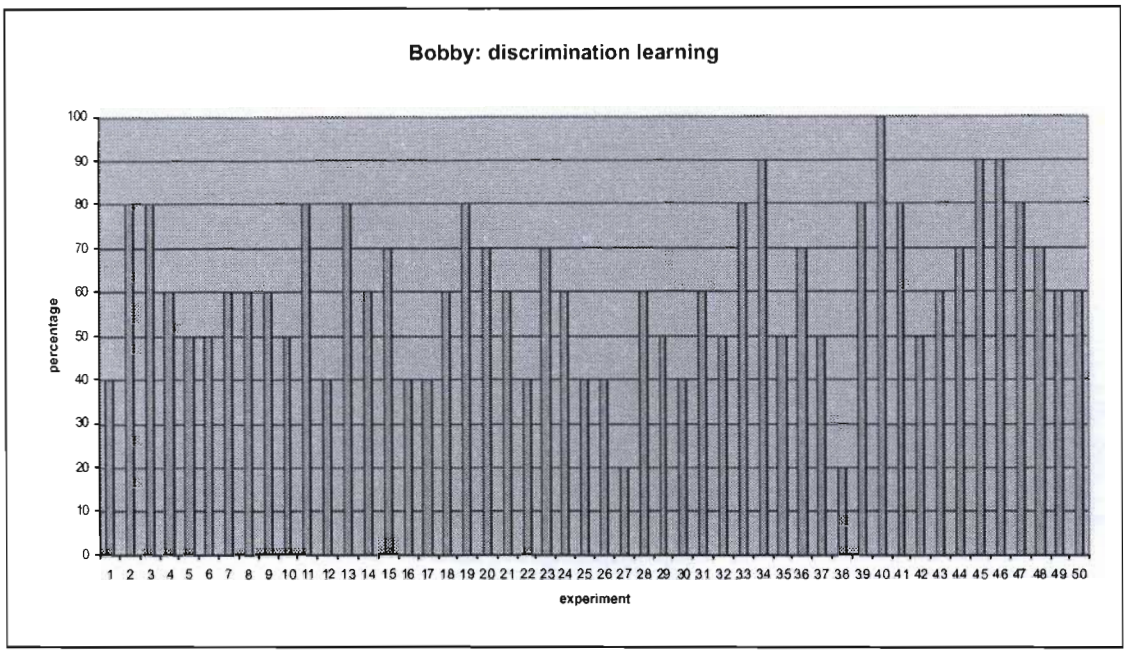

Figure 4.2: Percentage scores achieved by Bobby during discrimination learning experiments.

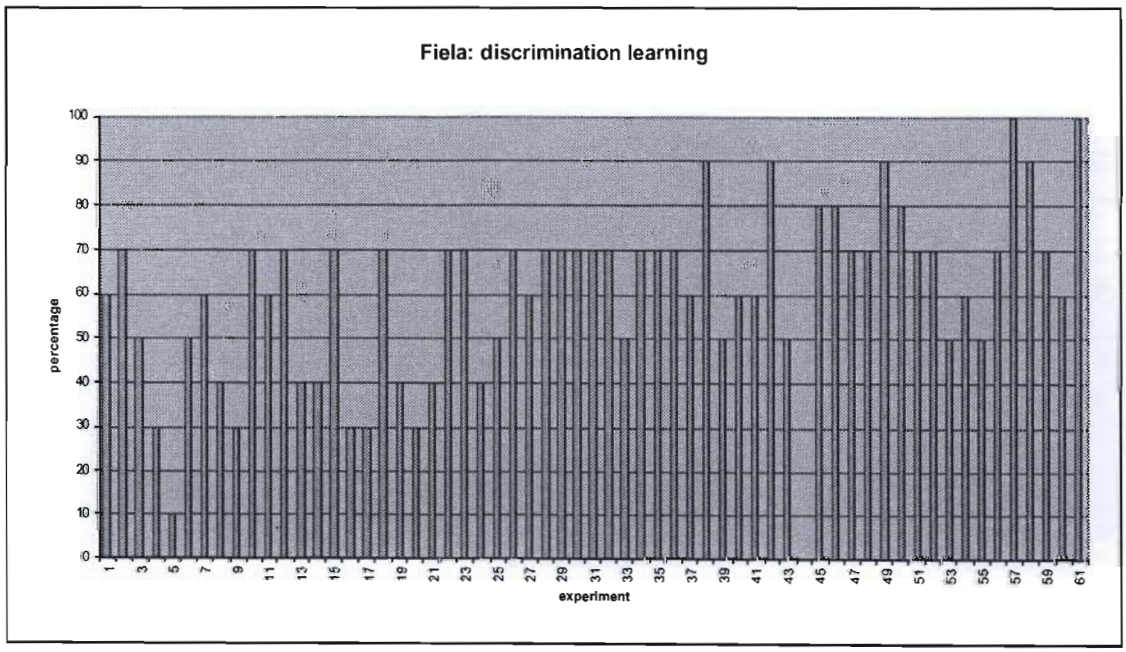

Figure 4.3: Percentage scores achieved by Fiela during discrimination learning experiments.

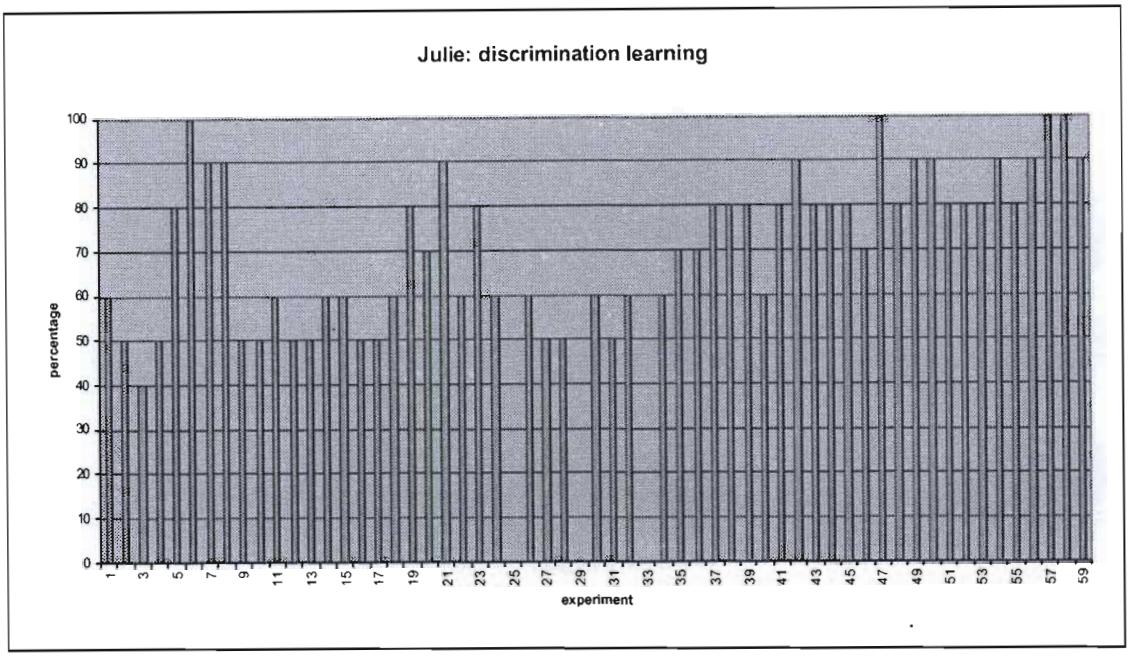

Figure 4.4: Percentage scores achieved by Julie during discrimination learning experiments.

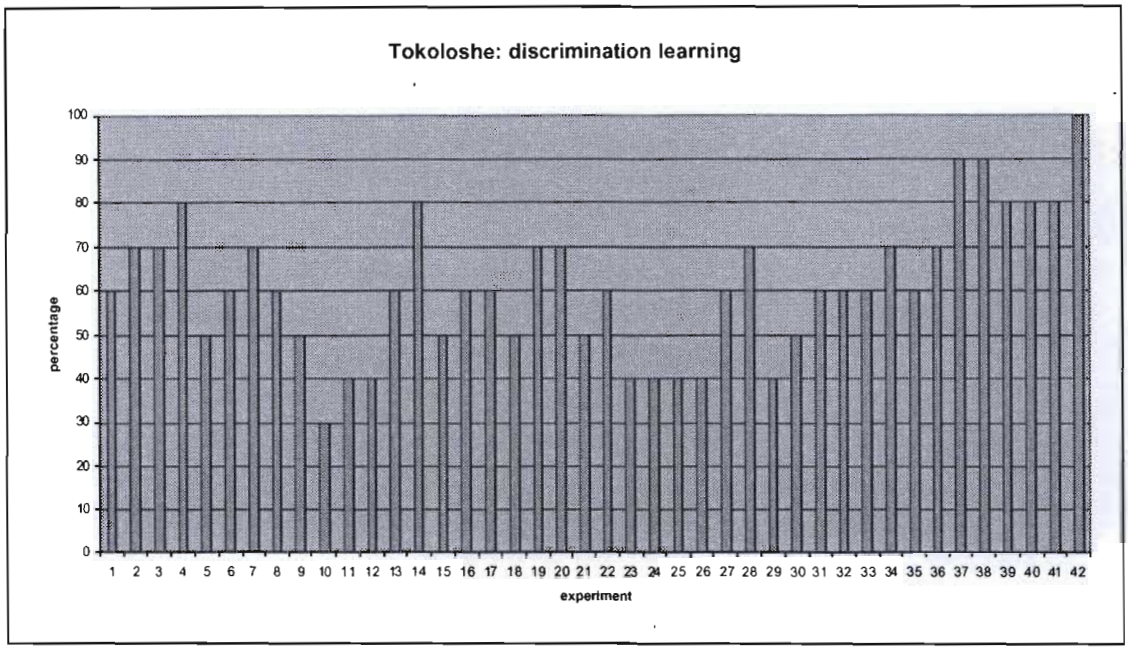

Figure 4.5: Percentage scores achieved by Tokoloshe during discrimination learning experiments.

Results from the discrimination task show a relatively even increase in scores through the course of the task. Bobby's (Fig.4.2) scores appear to fluctuate far more than the other subjects'. This is a result of her consistent response of opening the left hand side lid almost

exclusively throughout the trials. She later learnt to vary her response in relation to the stimulus. The low incidence of uncompleted experiments indicates the high level of motivation on all subjects, possibly due to the novelty value of the task and rewards at this stage. Julie's (Fig.4.4) higher incidence of uncompleted experiments is attributable to interruption of one experiment by feeding time, and another by a major fight in the vicinity. While Julie appears to have reached the criterion of 80% success over four experiments between experiments 5 and 8, these results were adjudged to have been as a result of sound cues, and not of discrimination between the stimuli. The experiments were therefore continued, with more care paid to the insertion of rewards to the apparatus. The immediate drop in her success rate may be attributed to this adjustment.

4.2 Discrimination – reversal

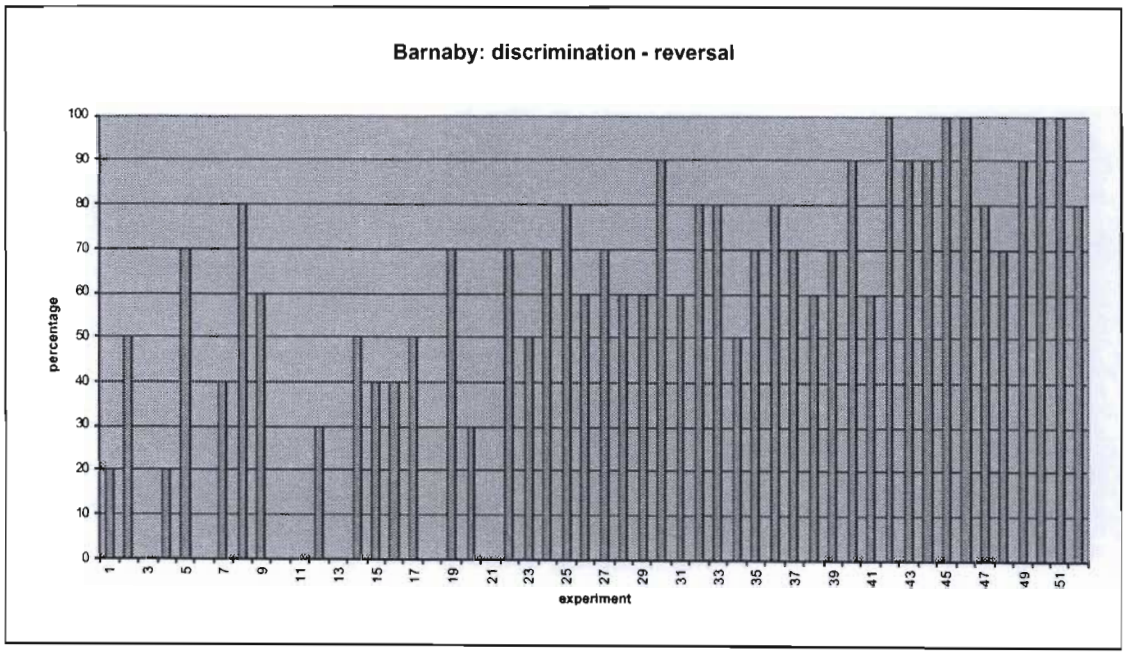

Figure 4.6: Percentage scores achieved by Barnaby during discrimination learning reversal experiments.

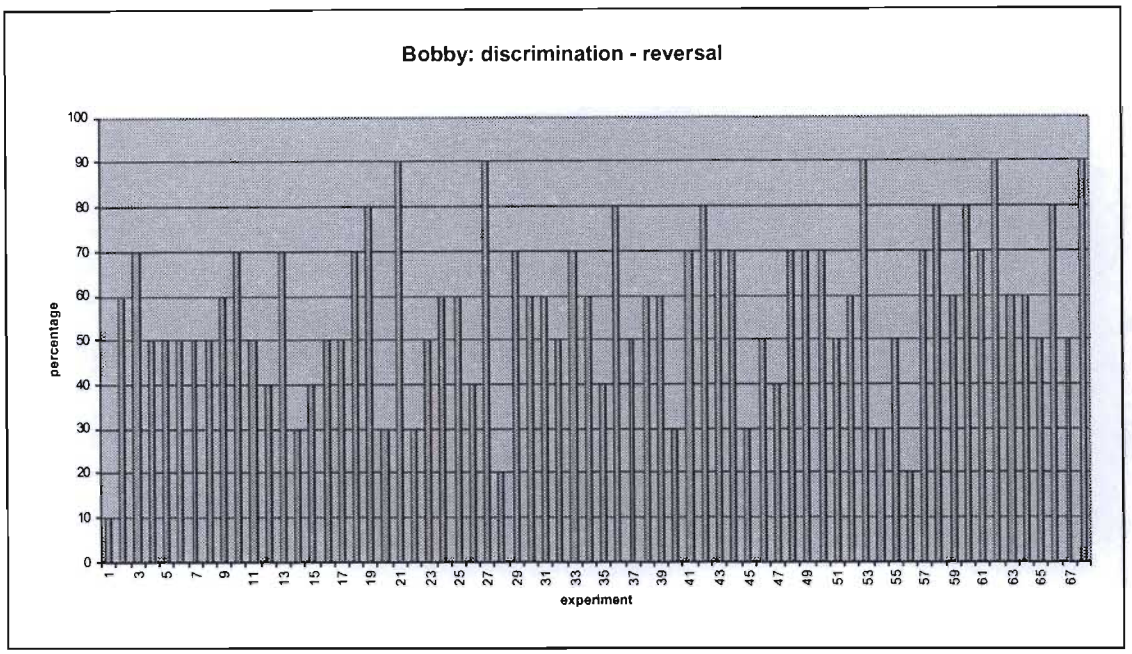

Figure 4.7: Percentage scores achieved by Bobby during discrimination learning reversal experiments.

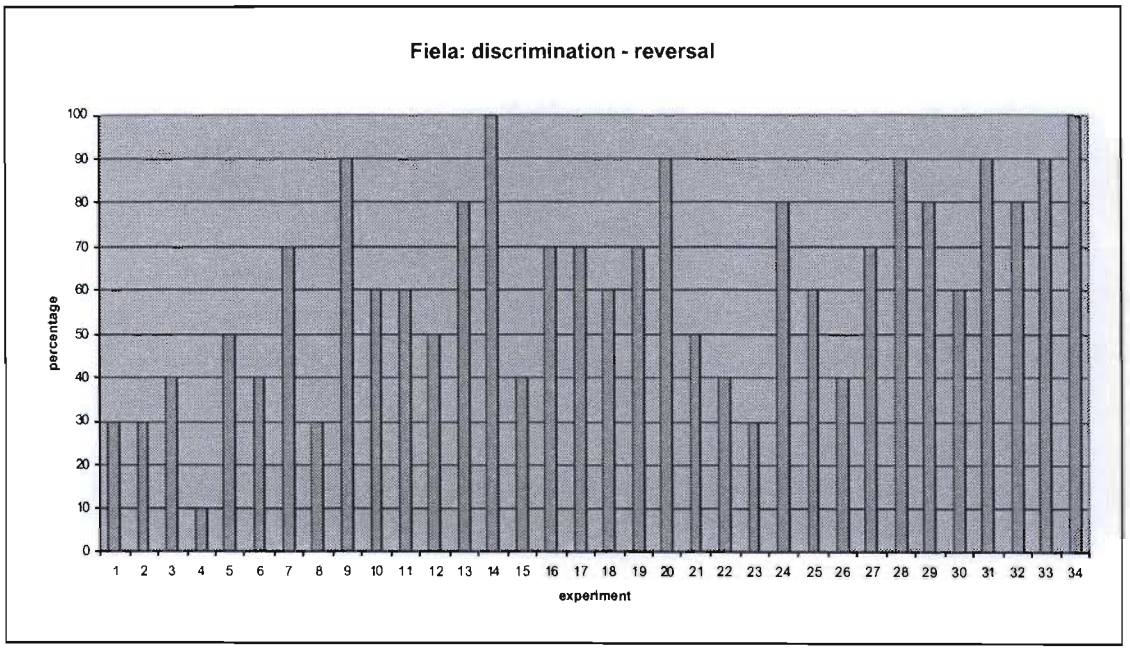

Figure 4.8: Percentage scores achieved by Fiela during discrimination learning reversal experiments.

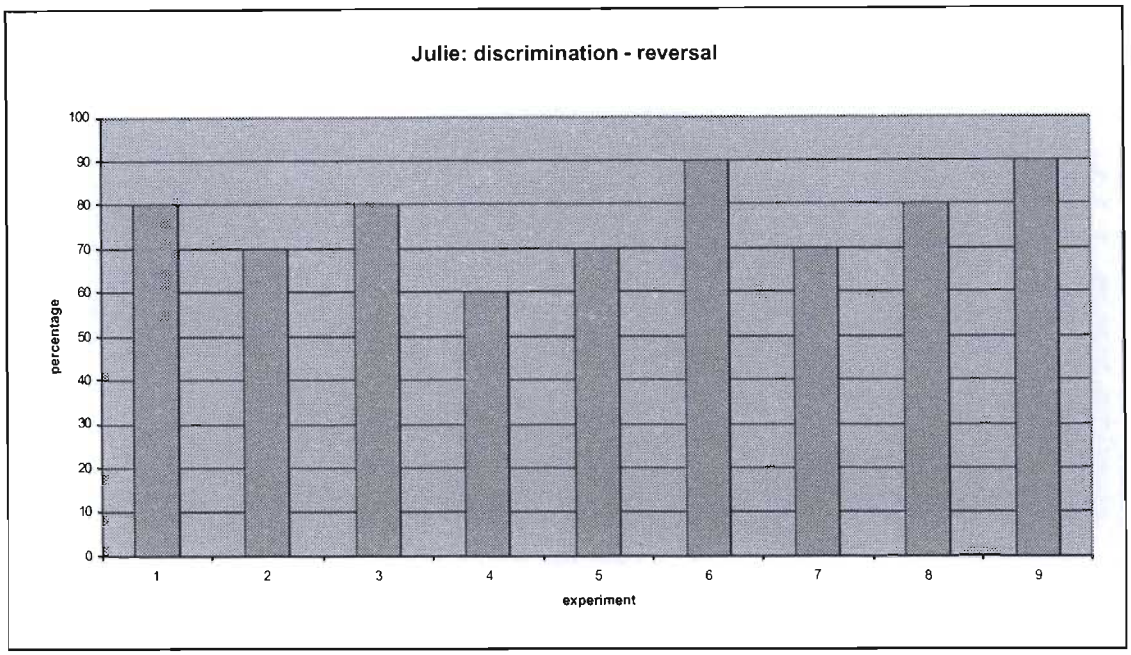

Figure 4.9: Percentage scores achieved by Julie during discrimination learning reversal experiments.

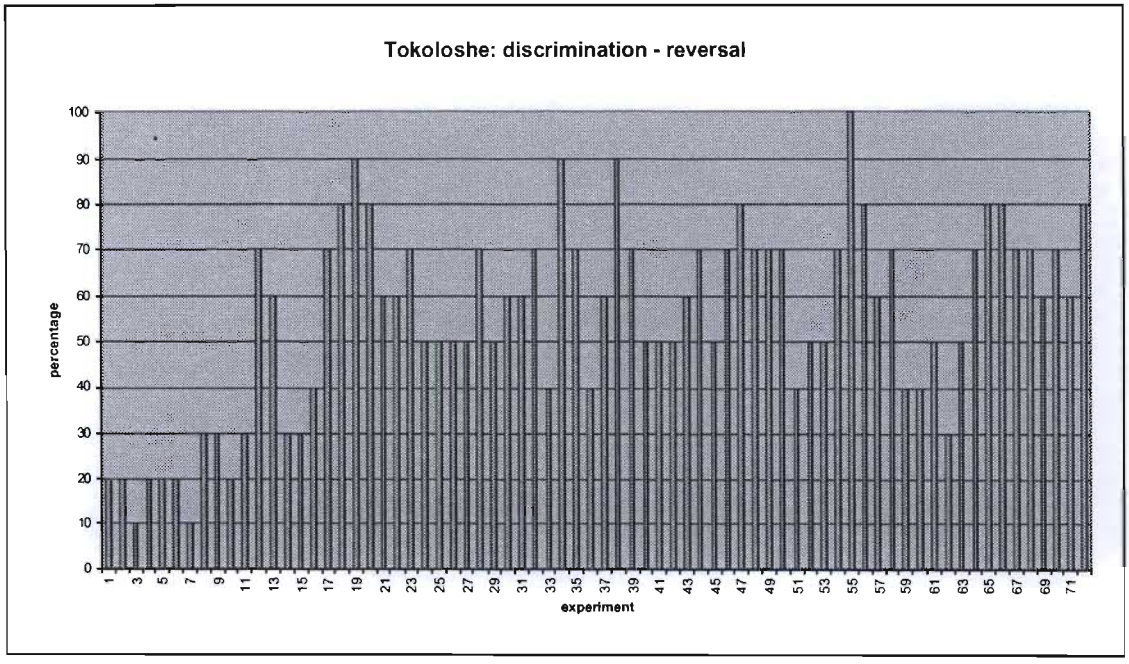

Figure 4.10: Percentage scores achieved by Tokoloshe during discrimination learning reversal experiments.

As with the discrimination results, the results for the discrimination reversal task showed a relatively even increase in scores over time, again with the exception of Bobby (Fig.4.7), as a

result of her positional bias. Barnaby (Fig.4.6), Tokoloshe (Fig.4.10), and Fiela (Fig.4.8) exhibited a rapid increase in scores almost to the point of criterion, but lacked the consistency of response to maintain it. As the experiments continued, their scores reached a plateau, probably as a result of boredom due to repetition. Julie (Fig.4.9) reached criterion easily in under ten experiments. Again, this may have been influenced by sound cues, as the rewards used at this time were peanuts – the loudest of the reward items in the apparatus. All the experiments were completed by the subjects, with the exception of Barnaby, who may have been simply uninterested by the rewards offered, or otherwise distracted by events around him, such as fights or females in oestrus.

4.3 Match-To-Sample (MTS)

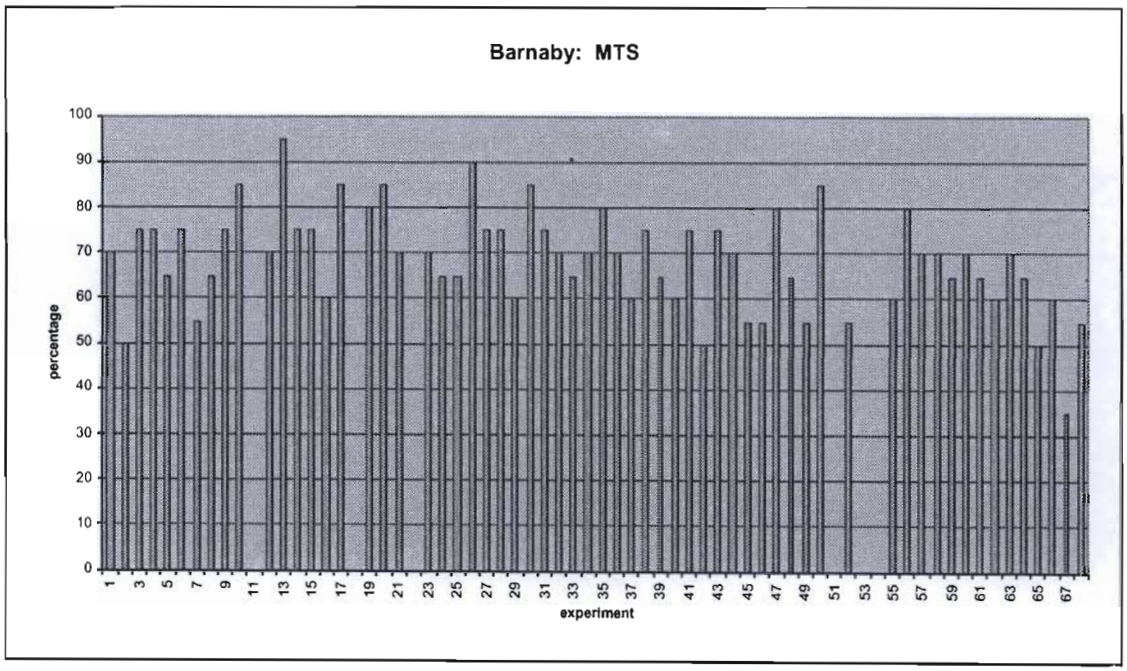

Figure 4.11: Percentage scores achieved by Barnaby during match-to-sample experiments.

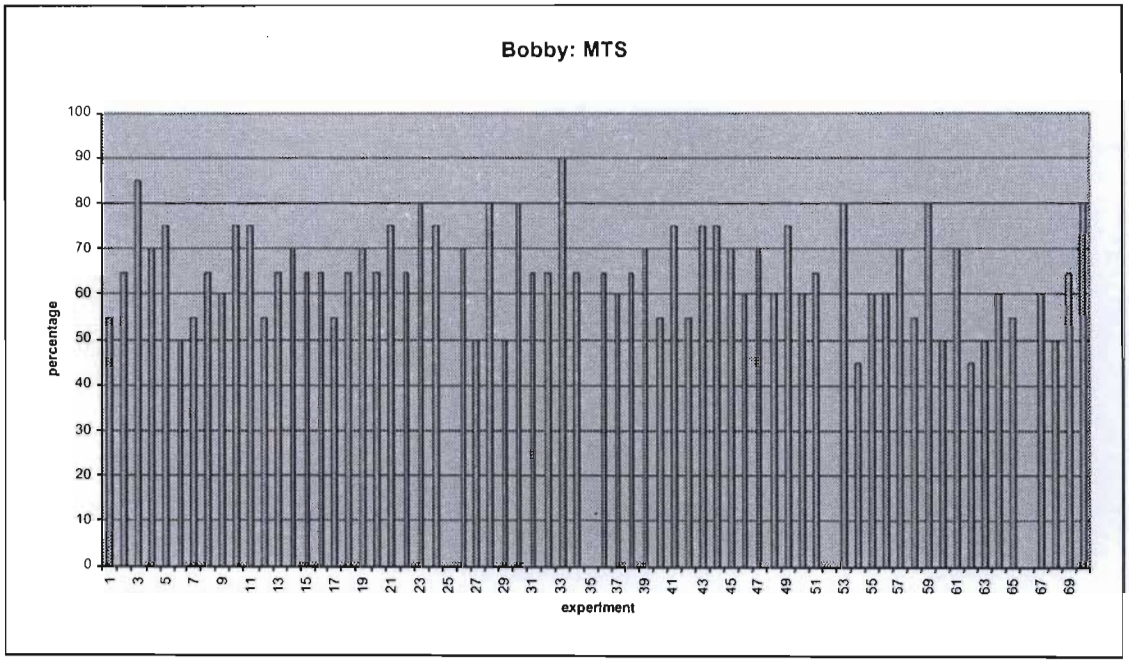

Figure 4.12: Percentage scores achieved by Bobby during match-to-sample experiments.

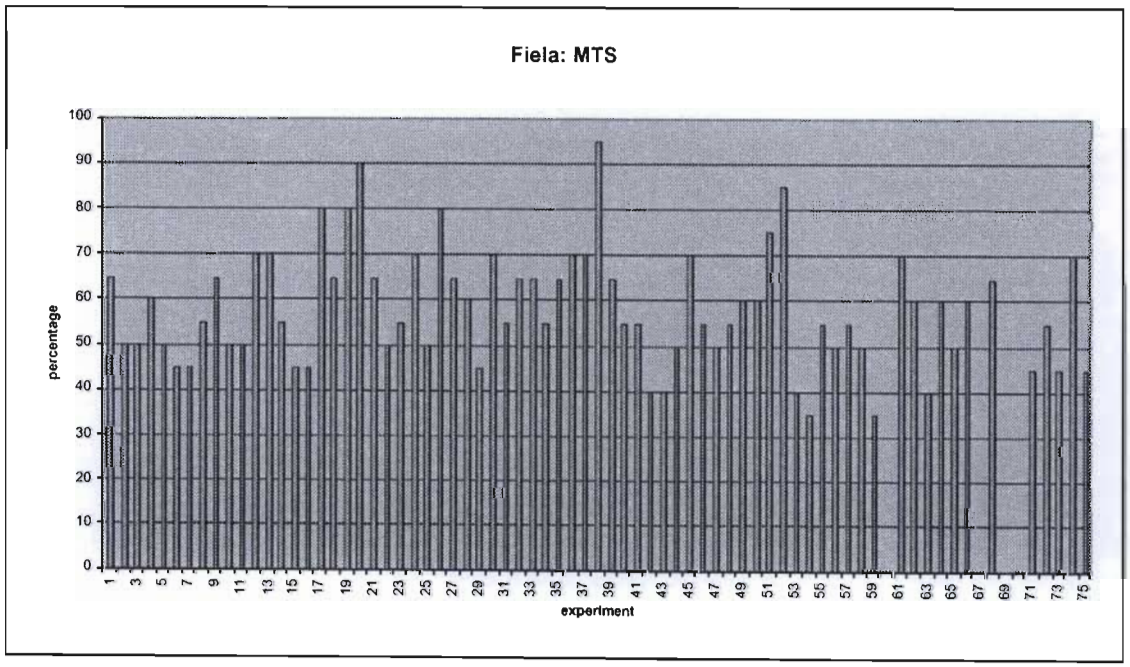

Figure 4.13: Percentage scores achieved by Fiela during match-to-sample experiments.

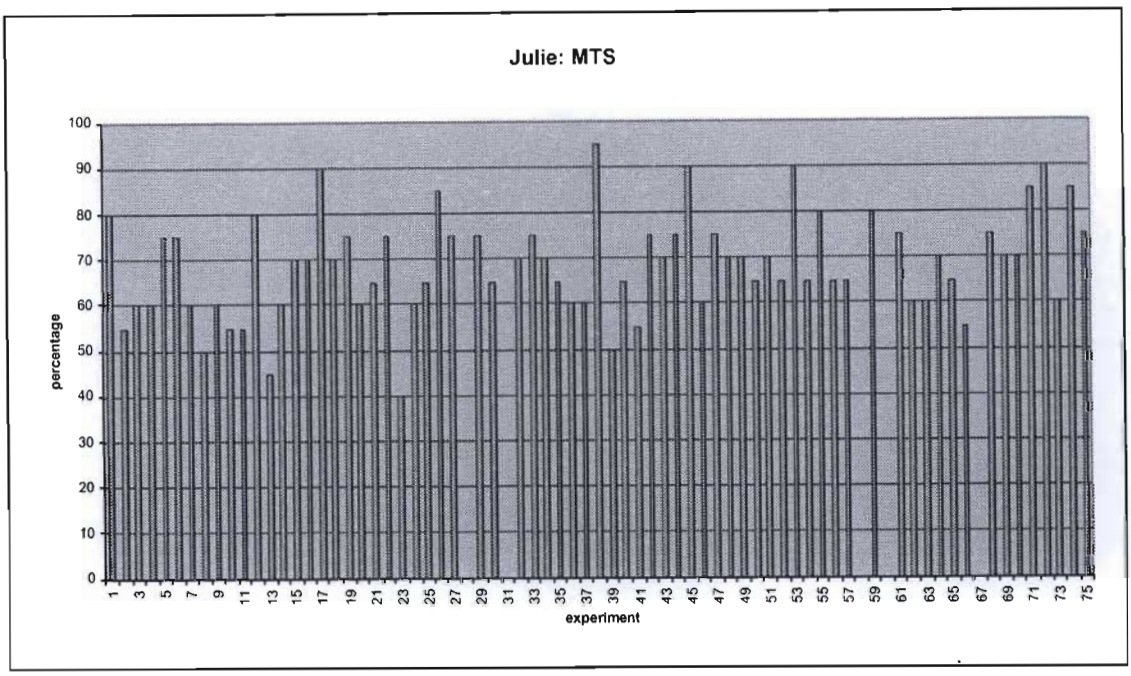

Figure 4.14: Percentage scores achieved by Julie during match-to-sample experiments.

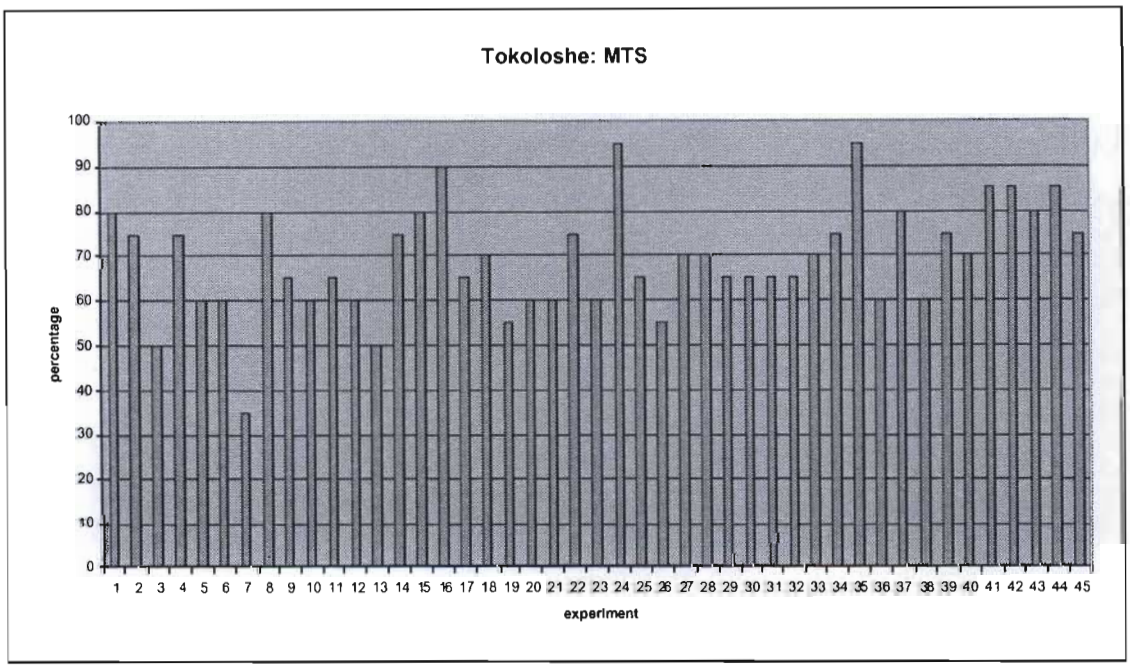

Figure 4.15: Percentage scores achieved by Tokoloshe during match-to-sample experiments.

With the match-to-sample task, it is evident there is an increase in the number of uncompleted experiments. This may be due to the increased complexity of the task, and the subsequent demands on the subjects' attention. The exception is Tokoloshe (Fig.4.15), and this is

attributable to her insatiable appetite for biltong, which appeared to keep her motivation from wavering. The steady increase in scores to criterion is also lacking in this task, with a far more fluctuating pattern emerging. In fact, scores appear to decrease with time, possibly indicating decreasing motivation due to boredom with the task, or dissatisfaction with the rewards. Only Tokoloshe and Julie reached criterion for this task.

4.4 Match-to-non-sample

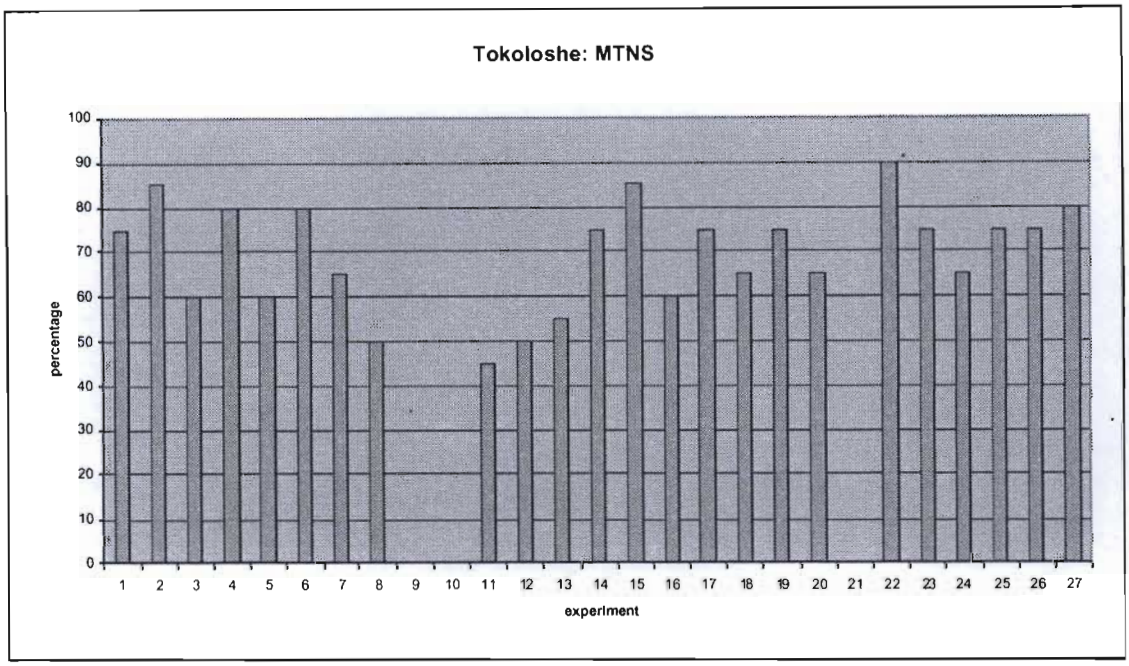

Figure 4.16: Percentage scores achieved by Tokoloshe during match-to-non-sample experiments.

As Tokoloshe (Fig.4.16) was the only subject to proceed to the match-to-non-sample task, only her results are presented here, and are not included for further analysis. The results show a gradual increase in score consistency towards criterion, though not reaching criterion due to the cessation of experimentation. The two successive uncompleted experiments (9 and 10) are as a result of almost total disinterest, even while using biltong as a reward. No further explanation can be ventured. The other uncompleted experiment (21) was a result of Tokoloshe breaking the apparatus so that no further trials were possible.

4.5 Comparison of tasks

The scores attained, times taken, and the number of distractions for each individual were analyzed separately using a one-way analysis of variance (ANOVA) for each different task.

For the discrimination task there was a significant difference between individuals' scores [$F_{4, 224} = 3.614$, $p<0.05$]. There appeared to be no significant differences between the times taken by individuals to perform these experiments [$F_{4, 224} = 1.201$, n.s], whereas there was a strongly significant difference between the distractions experienced by individuals [$F_{4, 224} = 6.959$, $p<0.01$] for the discrimination task.

Significant differences between individuals' times [$F_{4, 223} = 4{,}184$, $p<0.01$] and scores [$F_{4, 223} = 4{,}074$, $p<0.01$] were found for the discrimination reversal task, though no significant difference was found between individuals' distractibility [$F_{4, 223} = 2.226$, n.s].

In the match-to-sample task, significant differences were found for all of the variables. Distractibility [$F_{4, 309} = 13.968$, $p<0.001$], individual times [$F_{4, 309} = 4.465$, $p<0.01$], and individual scores [$F_{4, 309} = 10.760$, $p<0.001$] all showed significant differences between the individual subjects, with differences between both the individuals' distractibility and between their scores being highly significant.

The individuals' scores, times and distractions for each task were then combined. In keeping with the suggestion of Grafen and Hails (2002), these data were aggregated to avoid the complication of repeated measures, and then subjected to a one-way ANOVA to compare the overall differences between each task in each of the categories.

Neither the individuals' scores [$F_{2, 12} = 0.164$, n.s] nor the distractions [$F_{2, 12} = 1.388$, n.s], nor the times taken to complete the experiments [$F_{2, 12} = 0.410$, n.s] were found to vary significantly across the three different tasks.

4.6 Effects of distractions

A two-tailed Pearson correlation was performed on grouped data for each animal to explore the relationships between score and distraction, and time and distraction, for each subject, for each separate experiment. Only where there were fifty or more experiments completed by an individual for each task (giving at least five groups of ten experiments for analysis) was the data analysed. For this reason, Barnaby's discrimination results, Fiela and Julie's discrimination reversal results, and Tokoloshe's match-to-non-sample results were excluded from this analysis. Significant results are discussed here, with the full correlations being presented in Appendix 3.

There was a significant correlation between distractions and time for Barnaby in the discrimination reversal task ($r=0.968$, $p<0.01$), and no correlation between score and distraction for the same task. No correlations for Barnaby's match-to-sample task were discovered, either between score and distractions, or between time and distractions.

A negative correlation was found between score and distractions for Bobby's discrimination task ($r=-0.965$, $p<0.01$), though no correlation between distraction and time for the same task. No correlations resulted from Bobby's discrimination reversal results, though there was a correlation between distractions and time ($r=0.881$, $p<0.01$) for the match-to-sample task.

Fiela's discrimination results yielded a strong correlation between distractions and time ($r=0.901$, $p<0.01$), with no correlation between score and distractions. No correlations resulted from her match-to-sample task.

Julie's results for the discrimination task showed a correlation between distractions and time ($r=0.939$, $p<0.01$), and no correlation between score and distractions. No correlations resulted from her match-to-sample task.

There were no correlations resulting from analysis of Tokoloshe's discrimination and match-to-sample tasks. There was a correlation between distractions and time ($r=0.737$, $p<0.05$) for the discrimination reversal task, however.

For the individuals, the high number of correlations between distraction and time suggests that the time taken to complete an experiment corresponds with the number of distractions experienced. It seems that the distractions increase the time required. Similarly, the extremely low number of correlations between the score and the number of distractions suggests that the engaged social world the experiments are taking place in is not affecting the quality of the results. It has elsewhere been established that baboons are capable of performing these kinds of tasks (for example, Bovet and Vauclair, 1998; Depy *et al.*, 1999); it simply takes them longer to do so in this social environment.

Scatter graphs are used here to illustrate the relationships between the time taken to complete a set of ten experiments, and the number of distractions experienced by the subject for that set of experiments. Particularly in those tasks where a correlation between time and distractions was present, the almost linear relationship between the variables is most evident. Even in the tasks where there was no correlation found, the trend towards a linear relationship is sometimes evident.

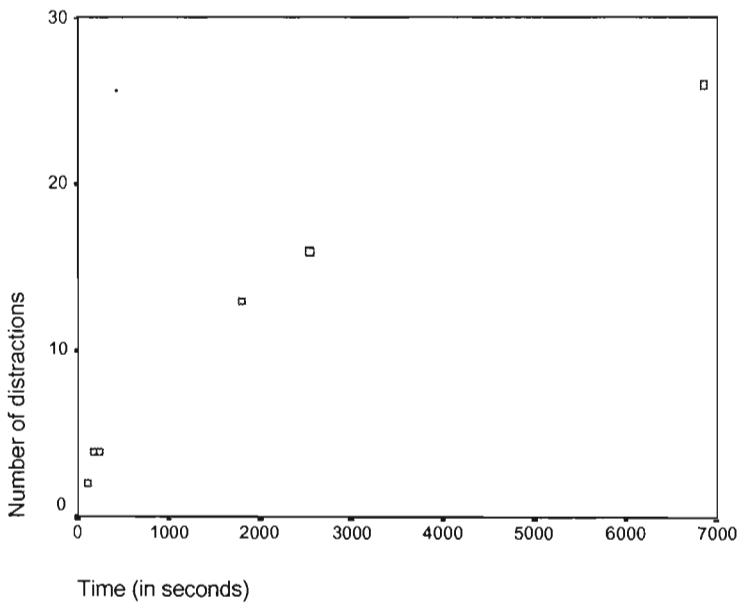

Figure 4.17: Number of distractions compared with time taken (in seconds) for Barnaby to complete discrimination reversal experiments ($r=0.968$, $p<0.01$)

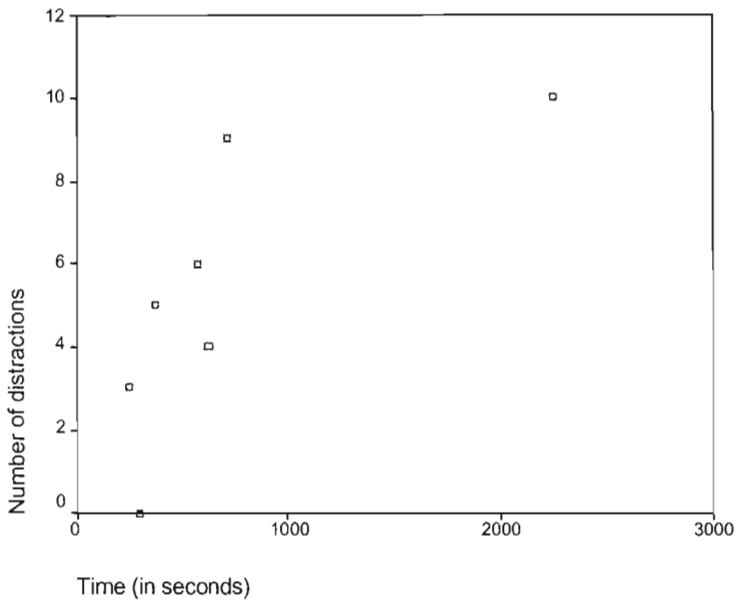

Figure 4.18: Number of distractions compared with time taken (in seconds) for Barnaby to complete match-to-sample experiments (r=0.739, n.s.)

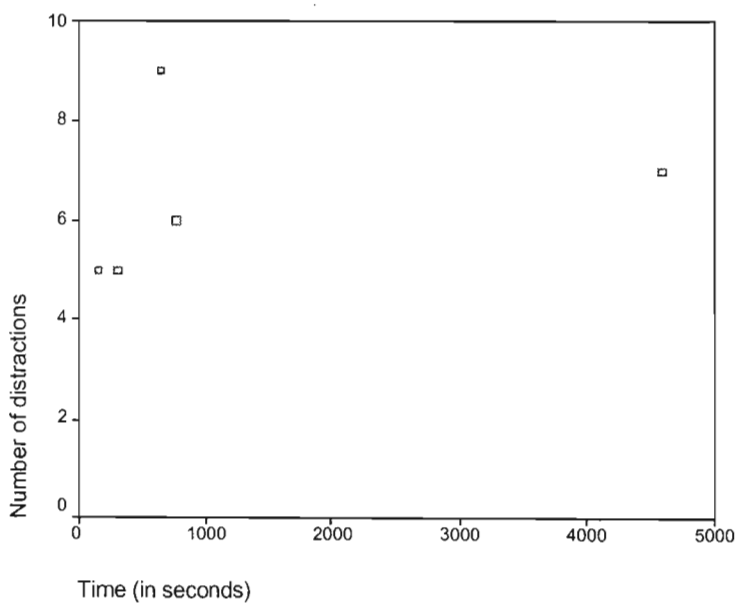

Figure 4.19: Number of distractions compared with time taken (in seconds) for Bobby to complete discrimination experiments (r=0.281, n.s.)

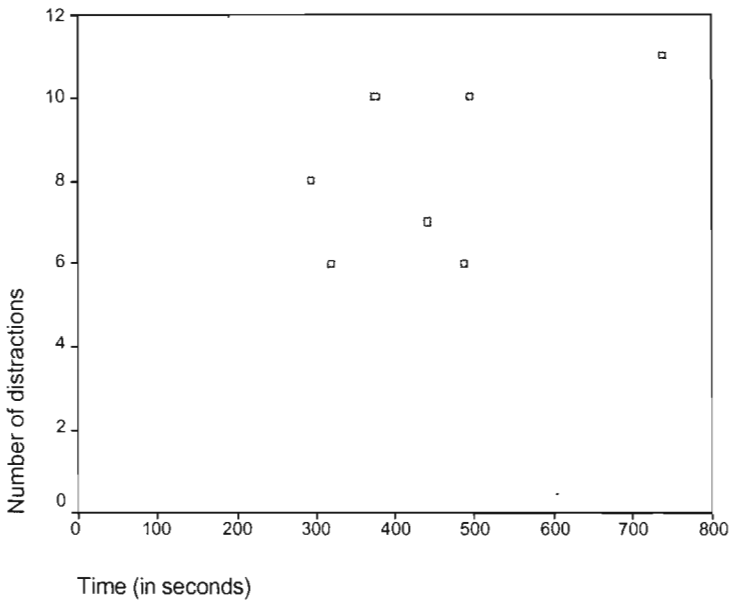

Figure 4.20: Number of distractions compared with time taken (in seconds) for Bobby to complete discrimination reversal experiments (r=0.545, n.s.)

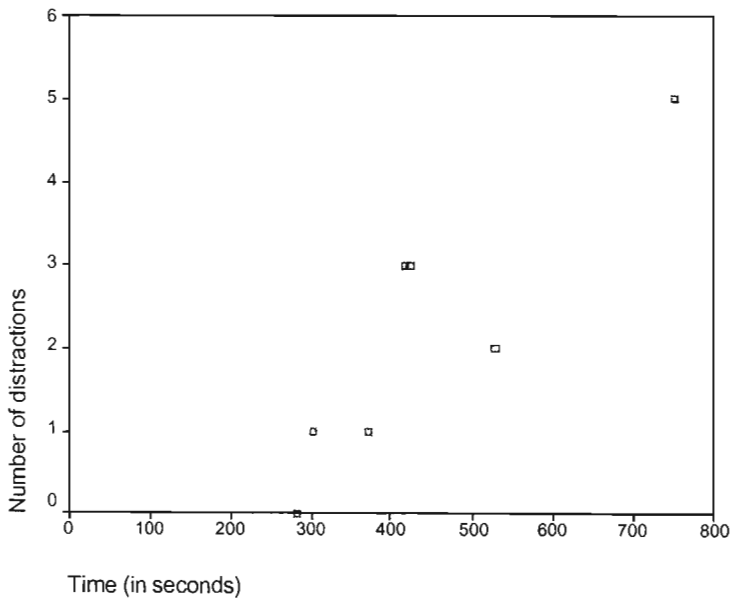

Figure 4.21: Number of distractions compared with time taken (in seconds) for Bobby to complete match-to-sample experiments (r=0.881, $p<0.01$)

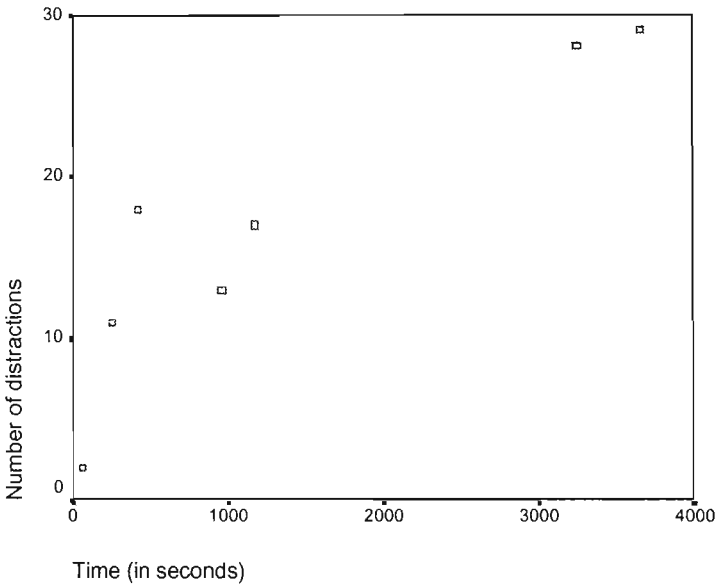

Figure 4.22: Number of distractions compared with time taken (in seconds) for Fiela to complete discrimination experiments ($r=0.901$, $p<0.01$)

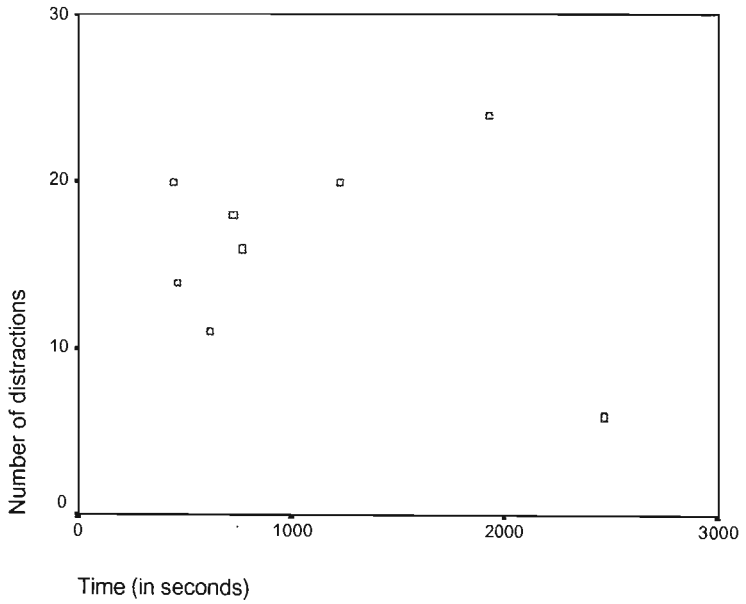

Figure 4.23: Number of distractions compared with time taken (in seconds) for Fiela to complete match-to-sample experiments ($r=-0.209$, n.s.)

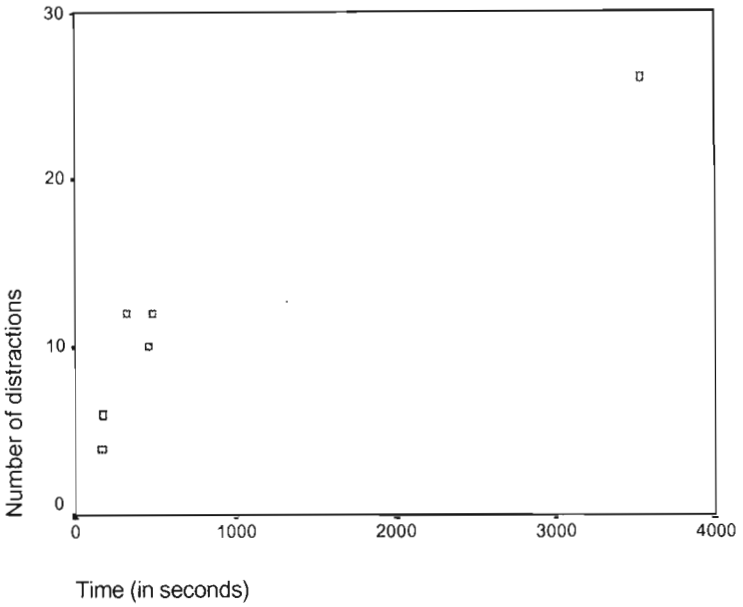

Figure 4.24: Number of distractions compared with time taken (in seconds) for Julie to complete discrimination experiments ($r=0.939$, $p<0.01$)

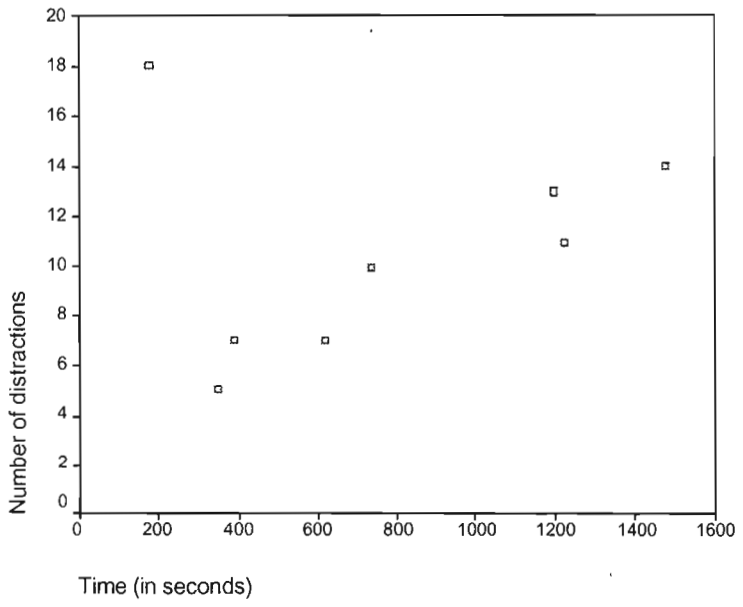

Figure 4.25: Number of distractions compared with time taken (in seconds) for Julie to complete match-to-sample experiments ($r=0.247$, n.s.)

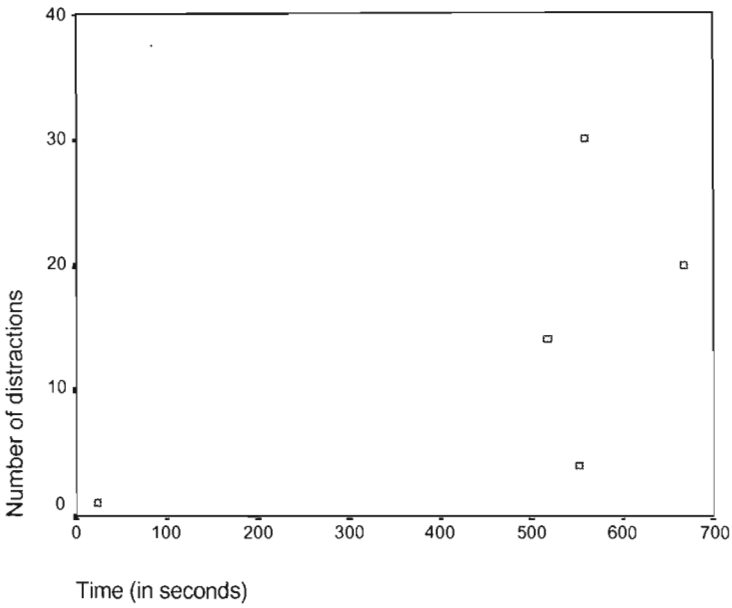

Figure 4.26: Number of distractions compared with time taken (in seconds) for Tokoloshe to complete discrimination experiments (r=0.634, n.s.)

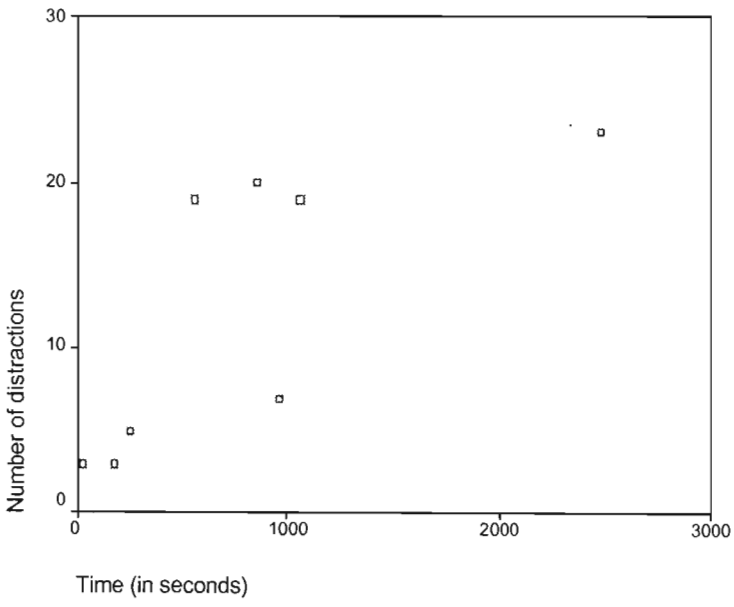

Figure 4.27: Number of distractions compared with time taken (in seconds) for Tokoloshe to complete discrimination reversal experiments (r=0.737, $p<0.05$)

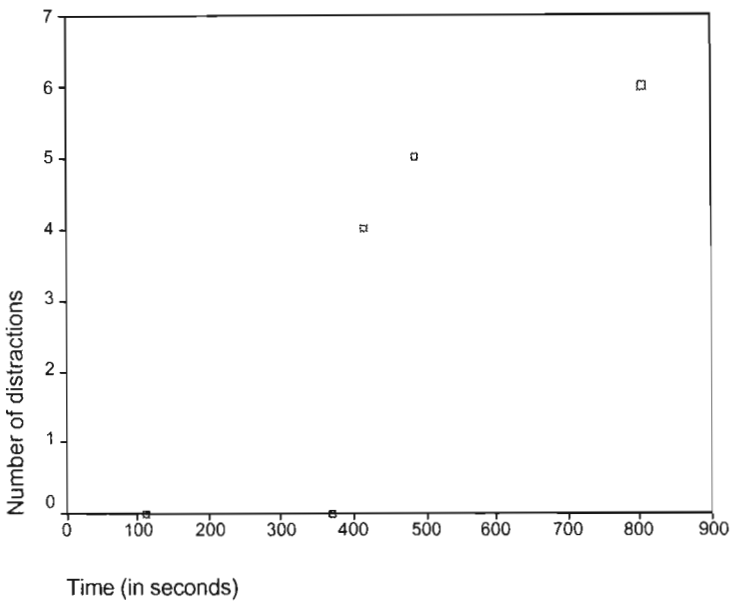

Figure 4.28: Number of distractions compared with time taken (in seconds) for Tokoloshe to complete match-to-sample experiments (r=0.836, n.s.)

5. DISCUSSION

The tasks presented in this study have tested the discrimination learning abilities of five socially housed chacma baboons, as well as their ability to discriminate between stimuli on the basis of perceptual cues in a match to sample task. The intended aim of investigating the analogical reasoning ability of these subjects was not achieved, due to experimental difficulties and time constraints. A number of studies that have used discriminations and match to sample tests to investigate aspects of analogical reasoning and concept formation (Depy *et al.*, 1999, Bovet and Vauclair, 1998) have also utilised basic discrimination tasks of a similar nature to this study as training tasks. These will be used as a basis of comparison for the effectiveness of the discrimination tasks of the present study.

In Bovet and Vauclair's (1998) initial training task, which required the subject to pull a rope to receive a food reward discriminating between two items (food and non-food), their subjects required a mean of 1460 trials to reach criterion of 80% success over 100 consecutive trials. In comparison, the subjects of this study required a mean of 466 trials (range 219-610) to reach criterion of 80% success over 40 consecutive trials (four consecutive experiments of ten trials). While these results are obviously not directly comparable due to the disparity in success criteria, it nevertheless gives an idea of the time taken to learn a discrimination task in these different experimental settings. The gradual learning of discriminative cues is consistent with Drea and Wallen's (1995) findings of gradual improvement in discrimination success over a series of trials with socially housed and tested rhesus monkeys.

The results of the discrimination reversal task confirm these findings. It would be expected that once the subjects had learned the initial discrimination task, their ability to transfer that learning to the reversed cues would be accelerated. Rather the converse is in order, as was also determined by Drea and Wallen (1995). The mean time to criterion was roughly the same as for the discrimination task, which suggests that the learning in these tasks is more likely a conditioned response rather than proof of any conceptual knowledge. For this reason it would have been especially relevant to proceed to the more complex analogical reasoning tasks to provide further evidence for or against baboons possessing abstract concept-forming capabilities. The only exception to this trend of gradual learning was Julie, who reached criterion in the discrimination reversal task in under 100 trials. This was probably as a result

of her perceptiveness to sound cues, and therefore cannot be considered to be reliable evidence of rapid learning.

The inconclusiveness of the match to sample results, and the variation of most of the subjects' scores, points towards issues of motivation or the increased effects of distractions on this more complex task. In Depy *et al.*'s (1999) match to sample training task in preparation for testing conceptual representation ability, baboons were able to learn to match samples using video joystick manipulation of a cursor to a criterion of 80% success over 120 trials in an average of 3024 trials. It is clear then that baboons are capable of performing these tasks. In the present study, only two individuals reached criterion of 80% success over 80 trials - Tokoloshe at 900 trials, and Julie at 1500 trials. Again the results are not directly comparable due to differences in success criteria, but the relative rapidity of learning response may be a result of the greater ecological validity of the present study, and the motivational effects of highly valued rewards.

A comparison of individuals' results between the three tasks showed no significant differences between the scores, number of distractions experienced, or the times taken to complete the experiments. This indicates that the complexity of the tasks to the individuals was similar, further indication of a conditioned learning response as opposed to conceptual recognition. This is contrary to the reports of other studies performed in this field, such as those of Martin-Malivel and Fagot (2001) and Drea and Wallen (1995), which recorded individual variation in response between tasks. As this study did not proceed to testing complex analogical reasoning tasks, these results are unsurprising.

To effect successful cognitive research under the conditions that exist at CARE, two conditions must be met. Either a task that requires only a little attention should be attempted, or the distractions should be limited – at least for the duration of the trial. Consequently this set of tasks is not as concerned with testing the actual cognitive performance of the subjects, but rather their ability to pay attention under conditions of multiple distractions.

Given these constraints, the fact emerges that experimental testing of cognitive processes is still possible, although the entire process may take substantially longer. The simplest explanation for this increase in experimental time is that these baboons live in a far more engaged social world, especially when compared to that of laboratory subjects, or those

removed from the social environment for the purposes of experimentation (eg. Bovet and Vauclair, 1998). Through the course of these tasks, it has also been proven that the experimental apparatus designed for this work is both practical and functional. While it is possible for the design to be further improved upon, the current design should prove more than adequate for future tasks of this nature.

The results clearly show the relationship between the number of distractions and the time taken to complete a set of trials. These results give credence to the idea that these individuals live in a more engaged social world, and as a result, the subjects' motivation wavers more than that of a laboratory subject. Motivation is also affected by the quality of reward on offer, as well as a number of other factors such as distractions and other environmental variables as discussed in Chapter 3. The lack of correlations between the number of distractions and the subjects' scores is encouraging, as it suggests that though it may take longer to complete an entire task, the results should not be adversely affected by the number of distractions. The quality of the results should be comparable to laboratory tests, only over a longer period of time.

6. CONCLUSION

The question of working with socially housed individuals is a complicated one. The subjects live in a far more engaged social world than their laboratory counterparts. Their motivation wavers and their distractibility is high, as they are well stimulated by the world around them and are not forced or coerced by manipulation of their needs to work on problems that are repetitive and probably boring.

It is possible to influence the motivation of individuals by means of increasing the variety and desirability of the rewards, and of the frequency of social contact. A protocol was devised to facilitate the application of discrimination and match-to-sample tasks, and their reversals, under these less-than-ideal conditions. An apparatus was constructed to suit the conditions, and proved to be functional and effective.

The tasks posed to the subjects proved within their capabilities, though over a longer time period than may be expected for the species as a result of the increased distractions and variable motivation levels. The scores achieved by the subjects for each experiment were not adversely affected by the number of distractions, though there were strong correlations between the time taken to complete an experiment and the number of distractions experienced. This suggests that results of experiments under these conditions should be of good quality, and comparable to that of laboratory subjects, considered over a longer time-period. While not all of the desired tasks could be performed as a result of time constraints on the collection of data, the groundwork has been laid for future experimental testing under similar conditions.

To optimise results under these conditions, a few changes could be made to the protocol used in this research, in conjunction with the improvements to the experimental design (see Appendix 2). Initially, attempts should be made to identify and ordinate the desirability of a wide range of food items for each subject. Once preferred items have been identified, they can be used to intersperse less favoured food items to provide variety and prevent motivation wavering to some extent.

The actual procedure of the trials is quite effective, and need not necessarily be altered. Time-out periods could be increased to two minutes, though at the risk of losing the attention of the subject to something else in that period. The increase in time would be useful, however, in providing a clear disincentive to an incorrect choice. The time-out period of one minute seemed insufficient, especially as the time taken to set up a new trial after the reward had been removed often approached 30 seconds, so there was no clear distinction between setting-up time, and time-out time.

Grooming of the experimenter would be allowed only after the full set of trials had been completed, to prevent inadvertently rewarding the subject while the experiment was in progress. By consistently allowing a grooming session at the end of an experiment, the motivation of the subject to complete the trials should be increased. However, this would only apply to those subjects that desired to groom the experimenter – in the case of this research, Fiela, Bobby, and Tokoloshe. Otherwise, some other reward – unique from the food rewards utilised in the trials – could be given to the subject to encourage motivation to work. For example, an egg could be given to the subject, as it is a highly desirable food item that is too large to use in the apparatus.

7. REFERENCES

Anderson, J.R; Awazu, S. & Fujita, K. (2000): Can squirrel monkeys (Saimiri sciureus) learn self-control? A study using food array selection tools and reverse reward contingency. *J.Exp. Psychol. Anim. B.* **26**(1): 87-97;

Baron-Cohen, S. (1990): Autism: a specific cognitive disorder of 'mind-blindness'. *Int. Rev. Psych.* **2**: 79-88;

Barrett, L; Dunbar, R.I.M. & Lycett, J.E. (2002): *Human Evolutionary Psychology.* Princeton University Press, Princeton;

Barton, R.A. & Dunbar, R.I.M. (1997): Evolution of the social brain. In: Whiten, A. & Byrne, R.W. (eds) *Machiavellian Intelligence II*, pp.240-263. Cambridge University Press, Cambridge;

Bechara, A; Damasio, A.R; Damasio, H; & Anderson, S.W. (1994): Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**: 7-12;

Boroditsky, L. (2000): Metaphoric structuring: understanding time through spatial metaphors. *Cognition* **75**:1-28;

Bovet, D. & Vauclair, J. (1998): Functional categorization of objects and of their pictures in baboons (Papio anubis). *Learn. Motiv.* **29**: 309-322;

Boysen, S.T; Berntson, G.G; Hannan, M.B; & Cacioppo, J.T. (1996): Quantity-based interference and symbolic representations in chimpanzees. *J. exp. Psych. Anim. Behav. Proc.* **22**: 76-86;

Brothers, L. (1997): *Friday's Footprint.* Oxford University Press, Oxford;

Burdyn, L.E. & Thomas, R.K. (1984): Conditional discrimination with conceptual simultaneous and successive cues in the squirrel monkey. *J. Comp. Psychol.* **98**: 405-413:

Byrne, R.W. (1995). *The Thinking Ape.* Oxford University Press, Oxford;

Byrne, R.W. & Whiten, A. (eds) (1988): *Machiavellian Intelligence.* Oxford University Press, Oxford;

Casperd, J. (1998): *Reconciliation among Chimpanzees*. PhD thesis, Liverpool University;

Cheney, D.L. & Seyfarth, R.M. (1990): *How Monkeys See the World.* Chicago University Press, Chicago;

Cheney, D.L; Seyfarth, R.M. & Silk, J.B. (1995): The role of grunts in reconciling opponents among baboons. *Anim. Behav.* **50**: 249-257;

Clayton, N.S. & Dickinson, A. (1998): Episodic-like memory during cache recovery by scrub jays. *Nature* **395**: 272-4;

Custance, D.M; Whiten, A; & Fredman, T. (2002): Social learning and primate reintroduction. *Int. J. Primat.* **23**: 479-499;

Dasser, V. (1987). A social concept in Java monkeys. *Anim. Behav.* **36**: 225-30;

Depy, D; Fagot, J. & Vauclair, J. (1999): Processing of above/below categorical spatial relations by baboons (Papio papio). *Behav. Processes* **48**: 1-9;

Donald, M. (1991): *Origins of the Modern Mind.* Harvard University Press, Cambridge MA;

Drea, C.M. & Wallen, K. (1995): Gradual acquisition of visual discrimination tasks in a social group of rhesus monkeys. *Anim. learn. & behav.* **23**: 1-8;

Dunbar, R.I.M. (1992): Neocortex size as a constraint on group size in primates. *J. human Evol.* **22**: 469-93.

Dunbar, R.I.M. (1998): The social brain hypothesis. *Evol. Anthropol.* **6**: 178-190;

Dunbar, R.I.M. (2000): Causal reasoning, mental rehearsal and the evolution of primate cognition. In: Heyes, C.M. & Huber, L. (eds) *Evolution of Cognition*, pp.205-231. MIT Press, Cambridge;

Dunbar, R.I.M. (2002): *Why are apes so smart?* In: Kappeler, P.M. & Pereira, M.E. (eds) *Primate Life Histories.* MIT Press, Cambridge;

Dunbar, R.I.M. & Bever, J. (1998): Neocortex size predicts group size in carnivores and some insectivores. *Ethology* **104**: 695-708;

Emery, N.J. & Perrett, D.I. (2000): How can studies of the monkey brain help us understand 'theory of mind' and autism in humans. In: Baron-Cohen, S; Tager-Flusberg, H. & Cohen, D.J. (eds) *Understanding other minds*, pp. 274-305. Oxford University Press, Oxford;

Grafen, A. & Hails, R. (2002): *Modern statistics for life scientists.* Oxford University Press, Oxford.

Hare, B. (2002): Can competitive paradigms increase the validity of experiments on primate social cognition? *Anim. Cogn.* **4**: 269-280;

Harris, P.L. (2000): *The Work of the Imagination.* Blackwell, Oxford;

Heyes, C.M. (1994): Reflections on self-recognition in primates. *Anim. Behav.* **47**: 909-19;

Heyes, C.M. (1998): Theory of mind in nonhuman primates. *Behav. Brain Sci.* **21**:101-34;

Kinderman, P; Dunbar, R.I.M. & Bentall, R.P. (1998): Theory-of-mind deficits and causal attributions. *Brit. J. Psych*. **89**: 191-204;

Limongelli, L; Boysen, S. & Visalberghi, E. (1995): Comprehension of cause-effect relationships in a tool-using task by chimpanzees. *J. comp. Psychol.* **109**: 18-26;

Martin-Malivel, J. and Fagot, J. (2001): Cross-modal integration and conceptual categorization in baboons. *Behav. Brain Res*. **122**: 209-213;

McPhail, E.M. (1985): Vertebrate intelligence: the null hypothesis. *Phil.Trans.R.Soc.Lond.* **308B**: 37-51;

Menzel, E.W. & Juno, C. (1984): Are learning sets learned? Or: Perhaps no nature-nurture issue has any simple answer. *Anim. learn. & behav.* **12**: 113-115;

O'Connell, S. (1995): *Theory of Mind in Chimpanzees.* PhD thesis, Liverpool University;

O'Connell, S. & Dunbar, R.I.M. (2002): *Comprehension of false belief: a comparative study of humans and chimpanzees*. (Unpublished manuscript);

Rapoport, S.I. (1999): How did the human brain evolve? *Brain Res. Bull*. **50**: 149-65;

Roth, D. & Leslie, A.M. (1998): Solving belief problems: toward a task analysis. *Cognition* **66**: 1-31;

Schrier, A.M. & Thompson, C.R. (1984): Are learning sets learned? A reply. *Anim. learn.& behav*. **12**: 109-112;

Semendeferi, K; Damasio, H; Frank, R. & van Hoesen, G.W. (1997): The evolution of the frontal lobes: a volumetric analysis based on three-dimensional reconstructions of magnetic resonance scans of human and ape brains. *J. human Evol.* **32**: 375-88;

Thompson, R.K.R. & Oden, D.L. (2000): Categorical perception and conceptual judgements by non-human primates: the paleological monkey and the analogical ape. *Cognitive Sci.* **24**: 363-396;

Tomasello, M. & Call, J. (1997): *Primate Cognition.* Oxford University Press, Oxford;

Tulving, E. (1983): *Elements of Episodic Memory*. Clarendon Press, Oxford;

Visalberghi, E. & Limongelli, L. (1994): Lack of comprehension of cause-effect relations in tool using capuchin monkeys. *J. comp. Psychol.* **108**: 15-22;

Whiten, A. (ed)(1991): *Natural Theories of Mind.* Oxford University Press, Oxford;

Whiten, A. & Byrne, R.W. (eds) (1997): *Machiavellian Intelligence II.* Cambridge University Press, Cambridge;

Wheeler, M.A. (1997): Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychol. Bull.* **121**: 331-354;

Wimmer, H. & Perner, J. (1983): Beliefs about beliefs. *Cognition* **13**: 103-28.

Internet sources

Http 1: http://nt.media.hku.hk/webcourse/references/tips/skinner.html (accessed 04-05-2004)

Http 2: http://www.phy.ilstu.edu/~wenning/ptefiles/311content/effective/animalhuman.html (accessed 04-05-2004)

8. APPENDIX 1 - SUBJECTS

Figure 8.1: Fiela

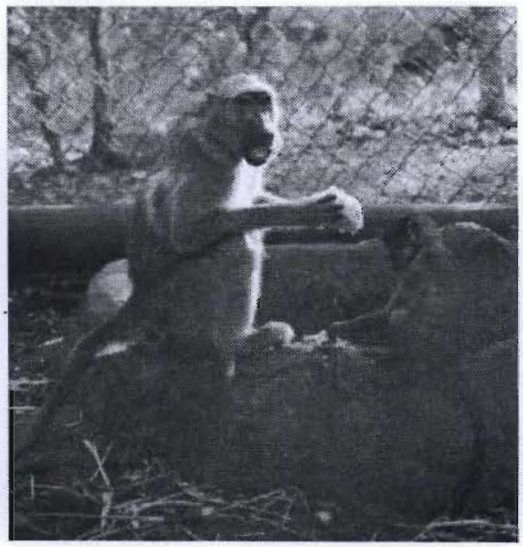

Figure 8.2: Julie (with her son, Wingnut)

Figure 8.3: Barnaby

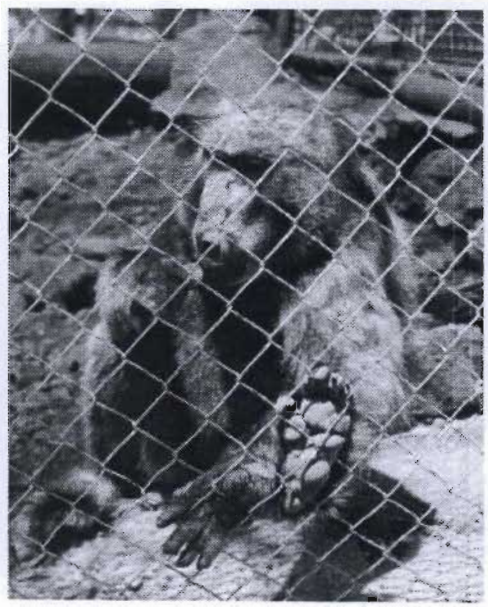

Figure 8.4: Bobby

Figure 8.5: Tokoloshe. Note multiple layers of wire.

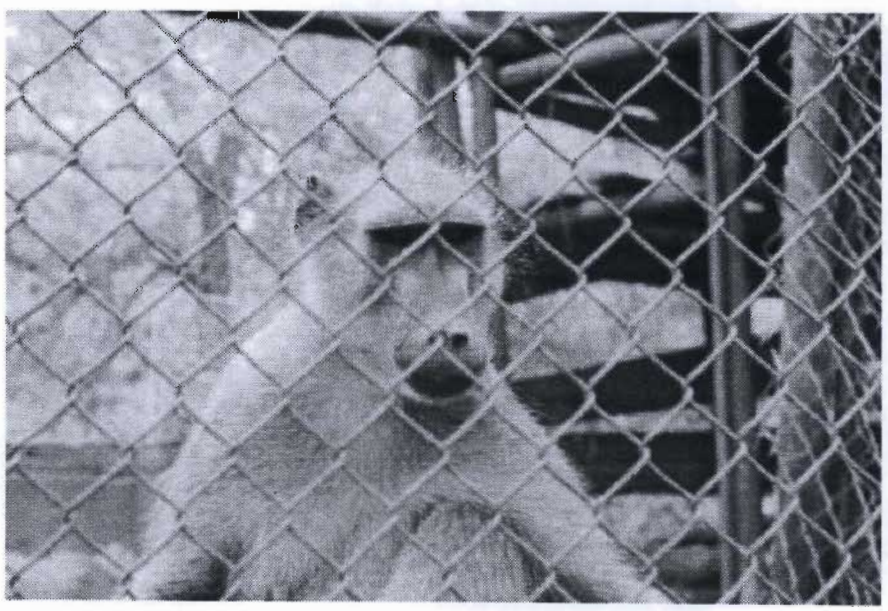

Figure 8.6: Tokoloshe, showing lightness of coat colouration.

9. APPENDIX 2 – APPARATUS PLAN

The dimensions for a functional apparatus are as follows:

Frame

Base: 60cm x 40cm masonite board

Side walls (x2): 2cm x 10cm x 38cm pine board

Back wall: 2cm x 10cm x 60cm pine board

Centre walls (x2): 2cm x 10cm x 38cm pine board

Brace struts (x2): 2cm x 10cm x 21cm pine board

Centre cover: 25cm x 10 cm masonite board

Lids

Base (x2): 20cm x 15cm masonite board

Back wall (x2): 2cm x 2cm x 15cm pine board

Side walls (x4): 2cm x 18cm strip masonite board

Overhangs (x4): 3cm x 18cm strip masonite board

Mountings

2x cup hooks

2x ring screws

2x 6cm hinges for flippers

4x 6cm hinges for lids

2x barrel bolts (flat sliding locks)

Flippers (x2): 2cm x 10cm x 21cm pine board

Tiles: 10.5cm x 17.5cm masonite board

Soundproofing

thick felt glued to interior of "box" areas including the walls, tacked down in the corners.

Suspension

2x 120cm parachute cord (2mm)

2x 6cm steel karabiners (snap-links)

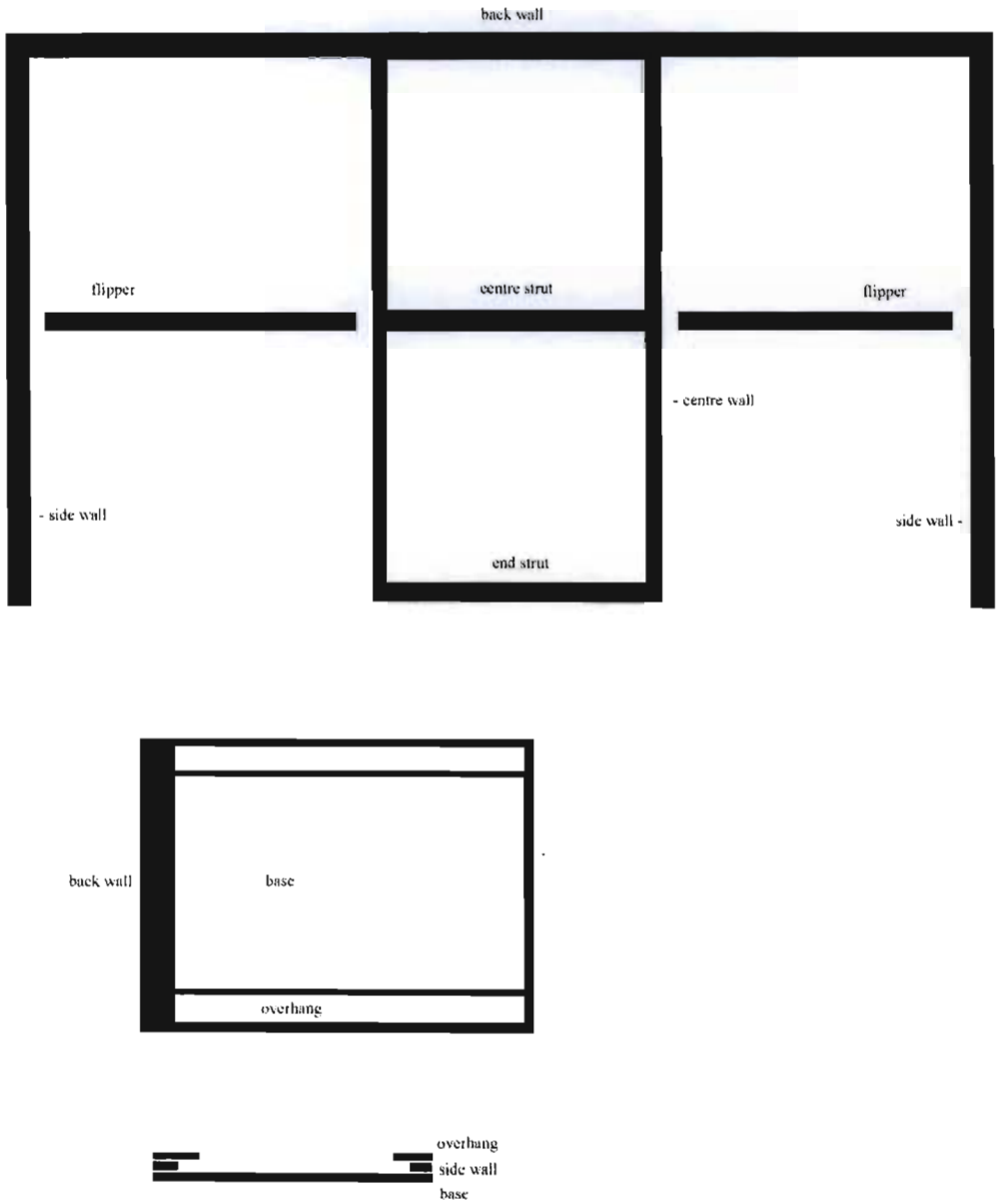

Figure 9.1: Plan diagram of apparatus design.

10. APPENDIX 3 – CORRELATIONS

Table 10.1: Correlations for Barnaby's discrimination reversal experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.277
	Sig. (2-tailed)	.	.596
	N	6	6
DISTRACT	Pearson Correlation	-.277	1.000
	Sig. (2-tailed)	.596	.
	N	6	6

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.968**
	Sig. (2-tailed)	.	.002
	N	6	6
TIME	Pearson Correlation	.968**	1.000
	Sig. (2-tailed)	.002	.
	N	6	6

**. Correlation is significant at the 0.01 level (2-tailed).

Table 10.2: Correlations for Barnaby's match to sample experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.243
	Sig. (2-tailed)	.	.599
	N	7	7
DISTRACT	Pearson Correlation	.243	1.000
	Sig. (2-tailed)	.599	.
	N	7	7

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.739
	Sig. (2-tailed)	.	.058
	N	7	7
TIME	Pearson Correlation	.739	1.000
	Sig. (2-tailed)	.058	.
	N	7	7

Table 10.3: Correlations for Bobby's discrimination experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.965**
	Sig. (2-tailed)	.	.008
	N	5	5
DISTRACT	Pearson Correlation	-.965**	1.000
	Sig. (2-tailed)	.008	.
	N	5	5

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.281
	Sig. (2-tailed)	.	.647
	N	5	5
TIME	Pearson Correlation	.281	1.000
	Sig. (2-tailed)	.647	.
	N	5	5

Table 10.4: Correlations for Bobby's discrimination reversal experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.233
	Sig. (2-tailed)	.	.616
	N	7	7
DISTRACT	Pearson Correlation	.233	1.000
	Sig. (2-tailed)	.616	.
	N	7	7

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.545
	Sig. (2-tailed)	.	.206
	N	7	7
TIME	Pearson Correlation	.545	1.000
	Sig. (2-tailed)	.206	.
	N	7	7

Table 10.5: Correlations for Bobby's match to sample experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.545
	Sig. (2-tailed)	.	.206
	N	7	7
DISTRACT	Pearson Correlation	.545	1.000
	Sig. (2-tailed)	.206	.
	N	7	7

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.881**
	Sig. (2-tailed)	.	.009
	N	7	7
TIME	Pearson Correlation	.881**	1.000
	Sig. (2-tailed)	.009	.
	N	7	7

**. Correlation is significant at the 0.01 level (2-tailed).

Table 10.6: Correlations for Fiela's discrimination experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.740
	Sig. (2-tailed)	.	.057
	N	7	7
DISTRACT	Pearson Correlation	.740	1.000
	Sig. (2-tailed)	.057	.
	N	7	7

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.901**
	Sig. (2-tailed)	.	.006
	N	7	7
TIME	Pearson Correlation	.901**	1.000
	Sig. (2-tailed)	.006	.
	N	7	7

**. Correlation is significant at the 0.01 level (2-tailed).

Table 10.7: Correlations for Fiela's match to sample experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.684
	Sig. (2-tailed)	.	.062
	N	8	8
DISTRACT	Pearson Correlation	-.684	1.000
	Sig. (2-tailed)	.062	.
	N	8	8

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	-.209
	Sig. (2-tailed)	.	.619
	N	8	8
TIME	Pearson Correlation	-.209	1.000
	Sig. (2-tailed)	.619	.
	N	8	8

Table 10.8: Correlations for Julie's discrimination experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.645
	Sig. (2-tailed)	.	.167
	N	6	6
DISTRACT	Pearson Correlation	-.645	1.000
	Sig. (2-tailed)	.167	.
	N	6	6

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.939**
	Sig. (2-tailed)	.	.005
	N	6	6
TIME	Pearson Correlation	.939**	1.000
	Sig. (2-tailed)	.005	.
	N	6	6

**. Correlation is significant at the 0.01 level (2-tailed).

Table 10.9: Correlations for Julie's match to sample experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.596
	Sig. (2-tailed)	.	.119
	N	8	8
DISTRACT	Pearson Correlation	-.596	1.000
	Sig. (2-tailed)	.119	.
	N	8	8

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.247
	Sig. (2-tailed)	.	.555
	N	8	8
TIME	Pearson Correlation	.247	1.000
	Sig. (2-tailed)	.555	.
	N	8	8

Table 10.10: Correlations for Tokoloshe's discrimination experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.598
	Sig. (2-tailed)	.	.287
	N	5	5
DISTRACT	Pearson Correlation	.598	1.000
	Sig. (2-tailed)	.287	.
	N	5	5

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.634
	Sig. (2-tailed)	.	.251
	N	5	5
TIME	Pearson Correlation	.634	1.000
	Sig. (2-tailed)	.251	.
	N	5	5

Table 10.11: Correlations for Tokoloshe's discrimination reversal experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	-.113
	Sig. (2-tailed)	.	.789
	N	8	8
DISTRACT	Pearson Correlation	-.113	1.000
	Sig. (2-tailed)	.789	.
	N	8	8

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.737*
	Sig. (2-tailed)	.	.037
	N	8	8
TIME	Pearson Correlation	.737*	1.000
	Sig. (2-tailed)	.037	.
	N	8	8

*. Correlation is significant at the 0.05 level (2-tailed).

Table 10.12: Correlations for Tokoloshe's match to sample experiments

Correlations

		SCORE	DISTRACT
SCORE	Pearson Correlation	1.000	.511
	Sig. (2-tailed)	.	.379
	N	5	5
DISTRACT	Pearson Correlation	.511	1.000
	Sig. (2-tailed)	.379	.
	N	5	5

Correlations

		DISTRACT	TIME
DISTRACT	Pearson Correlation	1.000	.836
	Sig. (2-tailed)	.	.078
	N	5	5
TIME	Pearson Correlation	.836	1.000
	Sig. (2-tailed)	.078	.
	N	5	5